

S t Nam Qu ry
side by side

Hit Count S t Nam
result set

DB=USPT,PGPB,JPAB,EPAB,DWPI; PLUR=YES; OP=OR

L6 fibrinogen adj2 binding and coagulase adj2 negative

3 L6

DB=USPT,PGPB,JPAB,EPAB,DWPI,TDBD; PLUR=YES; OP=OR

L5 fibrinogen adj2 binding and epidermidis

7 L5

DB=USPT,PGPB,JPAB,EPAB,DWPI; PLUR=YES; OP=OR

L4 (fbe or fig) adj5 gene\$ and staphylococc\$

0 L4

L3 (fbe or fig) adj5 gene\$ ans staphylococc\$

42895 L3

L2 (fbe or fig) adj5 gene\$

10 L2

L1 (fbe or fig) and epidermidis

0 L1

END OF SEARCH HISTORY

WEST[Help](#)[Logout](#)[Interrupt](#)[Main Menu](#)[Search Form](#)[Posting Counts](#)[Show S Numbers](#)[Edit S Numbers](#)[Preferences](#)[Cases](#)**Search Results -**

Terms	Documents
fibrinogen adj2 binding and coagulase adj2 negative	3

Database:

US Patents Full-Text Database
US Pre-Grant Publication Full-Text Database
JPO Abstracts Database
EPO Abstracts Database
Derwent World Patents Index
IBM Technical Disclosure Bulletins

S arch:

L6

[Refine Search](#)[Recall Text](#)[Clear](#)**Search History****DATE:** Tuesday, July 30, 2002 [Printable Copy](#) [Create Case](#)

WEST

Generate Collection

L3: Entry 20 of 25

File: EPAB

Oct 10, 1991

PUB-NO: DE003583987A1

DOCUMENT-IDENTIFIER: DE 3583987 A1

TITLE: TITLE DATA NOT AVAILABLE

PUBN-DATE: October 10, 1991

APPL-NO: DE03583987

APPL-DATE: October 31, 1985

PRIORITY-DATA: DE03583987A (October 31, 1985)

INT-CL (IPC): A61K 37/02; C07K 7/10; C12P 21/02

ABSTRACT:

A new antibiotic, designated epidermin (I), has the prim. structure: *Staphylococcus epidermidis* DSM 3095, which is resistant to (I), is new. (I) is made by aerobic cultivation of DSM 3095 at 34-37 deg.C on a complex nutrient soln. contg. 2-4% N source (e.g meat extract), 1-3% sugar or sugar alcohol; 0.25-1% alkaline earth carbonate and/or 0.25-0.5% alkaline earth hydroxide. The cells and inorganic salts are removed, then (I) isolated by (a) extracting with butanol at pH 8, evaporating the extract, dissolving the residue in MeOH and pptn. of lipids with ether or (b) adsorbing onto acrylic ester or polystyrene polymers, eluting with 99:1 MeOH-concn. H₂SO₄, neutralising with NH₃ and evaporating in vacuo. The isolate is then chromatographed on 'Sephadex LH-20'(RTM) to remove low mol.wt. peptides amino acid and salts, and subjected to liq-liq partitioning first in 3:1:3 n-butanol/ethyl acetate/0.1N acetic acid ((I) remaining at the starting position) and then in the neutral system 1:1 2-butanol/0.05N NH₄ acetate. Purified (I) is recovered as a colourless powder by freeze-drying.

elution with methanol/diluted hydrochloric acid, (c) the eluate is adjusted to a pH of 5.3 to 5.8, (d) the eluate is placed on a weak cation exchanger, (e) non-bound substances are subsequently washed out with a buffer solution at pH 7, (f) the active component is eluted out of the cation exchanger with a solution consisting of buffer substance, sodium chloride and methanol at pH 6.0 to 8.0 and for purification washed with water in order to remove salts and the epidermin is released from the resin with a methanol/acetic acid mixture and the solution is evaporated or freeze-dried, whilst the epidermin thus obtained may subsequently also be subjected to high performance liquid chromatography for extra purification.

CHOSEN-DRAWING: Dwg.0/10 Dwg.0/10

TITLE-TERMS: ISOLATE STAPHYLOCOCCUS CULTURE ADSORB POLYSTYRENE BASED
COPOLYMER ELUTION CHROMATOGRAPHY CATION EXCHANGE USEFUL ANTIBIOTIC TREAT
SKIN INFECTION

DERWENT-CLASS: A96 B04 D16

CPI-CODES: A04-B10; A04-C04; A12-M03; A12-V; A12-W11L; B02-E; B11-B; B11-C08D2; B12-A07;
D05-C02;

CHEMICAL-CODES:

Chemical Indexing M1 *01*

Fragmentation Code

H1 H100 H101 H181 H182 H4 H401 H481 H8 J0
J011 J012 J1 J171 J172 K0 L2 L250 M280 M311
M312 M313 M314 M315 M321 M331 M332 M333 M340 M342
M343 M349 M381 M391 M421 M510 M520 M530 M540 M620
M720 M903 N131 N161 Q233 V050 V901 V913 V923

Ring Index

63917

Registry Numbers

1327U 0502U

POLYMER-MULTIPUNCH-CODES-AND-KEY-SERIALS:

Key Serials: 0231 0306 3162 0418 1123 2020 2569 3264 3272 2769

Multipunch Codes: 014 034 04- 055 056 074 075 077 128 231 27& 473 53& 532 533
54& 623 624 642 645 720

SECONDARY-ACC-NO:

CPI Secondary Accession Numbers: C1990-006930

WEST



Generate Collection

L3: Entry 24 of 25

File: DWPI

Jun 10, 1997

DERWENT-ACC-NO: 1990-016158

DERWENT-WEEK: 199944

COPYRIGHT 2002 DERWENT INFORMATION LTD

TITLE: Isolating epidermin from staphylococcus epidermidis culture - by adsorption on styrene! based copolymer, elution and chromatography on cation exchanger, useful as antibiotic for treating skin infections

INVENTOR: FIEDLER, H; HOERNER, T; JUNG, G; KELLNER, R; WERNER, R; ZAEHNER, H;
FIEDLER, H P; HORNER, T; KELLNER, J R; WERNER, R G; ZAHNER, H

PATENT-ASSIGNEE:

ASSIGNEE

THOMAE GMBH KARL

CODE

THOM

PRIORITY-DATA: 1988US-0219698 (July 15, 1988)

PATENT-FAMILY:

PUB-NO	PUB-DATE	LANGUAGE	PAGES	MAIN-IPC
KR 9709289 B1	June 10, 1997		000	C07K001/14
EP 350810 A	January 17, 1990	G	018	
AU 8938103 A	January 18, 1990		000	
PT 91171 A	February 8, 1990		000	
DK 8903506 A	January 16, 1990		000	
JP 02084194 A	March 26, 1990		000	
ZA 8905362 A	March 27, 1991		000	
EP 350810 B1	September 29, 1993	G	021	C07K001/14
DE 58905744 G	November 4, 1993		000	C07K001/14
ES 2059645 T3	November 16, 1994		000	C07K001/14
IE 62402 B	January 25, 1995		000	C07K001/14
CA 1336896 C	September 5, 1995		000	C12P021/02
JP 2777205 B2	July 16, 1998		012	C12P021/02

DESIGNATED-STATES: AT BE CH DE ES FR GB GR IT LI LU NL SE AT BE CH DE ES FR GB GR IT LI
LU NL SE

CITED-DOCUMENTS:1.Jnl.Ref; A3...9139 ; EP 181578 ; EP 27710 ; No-SR.Pub

APPLICATION-DATA:

PUB-NO	APPL-DATE	APPL-NO	DESCRIPTOR
KR 9709289B1	July 15, 1989	1989KR-0010095	
EP 350810A	July 7, 1989	1989EP-0112446	
JP02084194A	July 14, 1989	1989JP-0182372	
ZA 8905362A	July 14, 1989	1989ZA-0005362	
EP 350810B1	July 7, 1989	1989EP-0112446	
DE58905744G	July 7, 1989	1989DE-0505744	
DE58905744G	July 7, 1989	1989EP-0112446	
DE58905744G		EP 350810	Based on
ES 2059645T3	July 7, 1989	1989EP-0112446	
ES 2059645T3		EP 350810	Based on
IE 62402B	July 14, 1989	1989IE-0002283	
CA 1336896C	July 14, 1989	1989CA-0605673	
JP 2777205B2	July 14, 1989	1989JP-0182372	
JP 2777205B2		JP 2084194	Previous Publ.

INT-CL (IPC): A61K 0/00; C07G 11/00; C07K 1/14; C07K 3/12; C07K 7/10; C07K 15/04; C07K 17/10; C12N 11/08; C12P 1/04; C12P 21/02; C12R 1/45; C12P 21/02; C12R 1/45; C12P 21/02; C12R 1/45

ABSTRACTED-PUB-NO: EP 350810A

BASIC-ABSTRACT:

the polypeptide antibiotic epidermin (I) is isolated and purified from a culture of a *Staphylococcus epidermidis* strain by (a) applying the culture broth or filtrate to a styrene-divinyl copolymer (A), (2) eluting active ingredients with MeOH-dil HCl, (3) adjusting eluate to pH 5.3-5.8; (4) applying to a weak cation exchanger (B) (5) washing-out non-bound cpds. with pH7 buffer, (6) eluting (I) with pH 6-8 buffer contg. and MeOH, (7) readsorbing (I) onto (A) washing the resin with water (desalting) and eluting with MeOH-MeCOOH mitd., (8) evaporating or freeze-drying the eluate, and opt. (9) further purifying by h.p.l.c.

The *S. epidermidis* strains used are pref. DSM 3095 or NC18 11536.

USE/ADVANTAGE - (I) is known for treatment of skin infections such as eczema, impetigo, cellulitis and acne. This method is simple and produces significantly higher yields of (I) then known processes.

ABSTRACTED-PUB-NO:

EP 350810B

EQUIVALENT-ABSTRACTS:

Process for isolating epidermin from a culture broth or a culture filtrate of a strain of *Staphylococcus epidermidis* and for purifying this substrate, characterised in that (a) the culture filtrate or culture broth is added to a styrene-divinyl copolymer, (b) the active component is released from the resin by

A:Reference number: S41539; MUID:94224142
 A:Accession: S41539
 A>Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1933 <MCD>
 A:Cross-references: EMBL:Z18852; NID:9397525; PIDN:CAA79304.1; PID:9397526

Query Match 14.6%; Score 452; DB 2; Length 933;
 Best Local Similarity 26.6%; Pred. No. 3; ie-12;
 Matches 154; Conservative 105; Mismatches 234; Indels 86; Gaps 26;

```

QY 8 SDEKNDVNNQISINTDNNQIIRKKEETNNYDGIKRSEDRTSTNVDENAFLQK 67
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 51 SNEKSNDSVSAAPRTDITN-----SDTKSSNTN---NGETSVAO 91
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 68 TPQDNLTEEEVKESSSVESNSSIDTAQPSHTTINREESVOTSDNVDSVDFANS 127
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 92 NP-----AOEQTQSSSTNATTEETPVTEGATTTTNOANTPATQDS-SNTNAELVN- 143
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 128 KIKESNDSKKEENTIDQPNKVEDSTTSOPSGYTNIDKISNODELLN-LPINEYENLA-186
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 144 --QTSNETTFNDITV-----SSVNSPQNTNANENSTODTTEATPSN--NES 189
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 187 RPLSTTSAPSIKRYTVN-----QLAAEAGSNVNHILKYTDQ--SITEGYDSEGV 235
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 190 APQSTDAKSNKVVNAVNTSAPRMRAPSLAAVADAPAGDITNQLNTVYGI-DSGTT 248
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 236 IKANDAEMLIYDVFEVDKAKSGDTMTVIDKNTVSQDLTSTTIRKIDNSELITAG 295
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 249 VYRQAGYVAKLYGFSVPNSAVKGDTEKITVPKELNLNGVTSTAKVPPIMAGD-QVLANG 307
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 296 TYDKNKQITTFPDYDYDKENIKAKHLKLTYSIDKSKVPNNNTKL-DVEYKTALS--VN 352
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 308 VIDS-DGNVYITFTDYNTKDVAKATLMPAYID-----PEVKKTKGVTLATIGSTTAN 362
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 353 KTIYEVORPENRTANLQSMETNIDRKNTVEQTIYINPL--RYSAKETNVNISGNGDE 410
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 363 KTVLDYDEKYEKRYVNLISGTTIDQIDKTNNTYRQTIYVNPFGDNVIAPVLTGKPKPTDS 422
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 411 GSTIIDSTIIKRYKVGONLPPDSNRIDYSEVEDYTNNDYDALGNNNDVINFG---- 466
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 423 NALIDQNTSIRKVVADNAADLSEYFV-NPENEDVNTSNITFPNPNQYKVEFNTPPD 481
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 467 NIDSPYIKVYSKYDPN-KDDYTTIQOTVMTQTTINETYGE--FRISYDNTIAFSTSSG 523
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 482 OTTPIYIYVNGHIDPNKGD-----LALRSTLYGYNSNIIMRSMWDNEVAFNNGSG 534
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 524 QGQG-DLP--PEKTYKIDY--VWEDVDKD-GIONTNDN 556
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 535 SGGDIDKPVVPEQDPDEPEIEPIEDSDSDPSDGSDDS 573
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

```

Query Match 15.9%; Score 490; DB 2; Length 953;
 Best Local Similarity 28.2%; Pred. No. 7; 7e-14;
 Matches 162; Conservative 83; Mismatches 236; Indels 94; Gaps 22;

```

QY 42 GIEKRSDETESTTVDNEATFLQKTPQDNLHTEEEVKESSSVESNSSIDTAQPSH 101
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 45 GHEAKAAEHTNGLNOSKNETT---APSEMK--TTEKY--DSROLKNDTQATADQPKV 96
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 102 TTINREESVOTSDNVDSVDFANSKIKESNDSKKEENTIDQPNKVEDSTTSOPSGY 161
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 97 T-----MDSATVKEKTSNMOS-----PQNTASOSTTQTSNV 129
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 162 TNIDEK---ISNODELLNLPINEYEN-KARPLSTTSAPSIKRYTVNQLAA-EOGSNVNH 216
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 130 TTNDKSSSTTYSNETDKSNL--TOAKNVSTTPKTTIKQALNKAENVTVAAAPQSTNVND 187
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 217 LKVTDOSI-----TEGYDSEGYKAKDAENLIYDVFEVDKVKSGDTMT 263
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 188 KYHFTNIDIALDKGVNKTGTGTEFWATSSDVLK-----LKANYTIDDSVKEGDTFT 239
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 264 VIDKNTVPSDLTDSFTTPIKIDNSELITAGTYDNKKQITTFPDYDYDKENIKAKHLK 323
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 240 FRYGGYFRPGSVRLPSQTONLNAQGNIIAKGIYDSTKTTTNTTNTVDDQTTNVSQSP 299
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 324 LFSYIDKSVPPNNNTKLDEYKTAALSVKNTTVEYORPENRTANLQSMETNIDTKNHT 383
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 300 QYAFKAKRENATDKTAIKKHEVTLGNDYTSKQIVDY---GNKRGQQLLSSTYINNEELS 356
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 384 VEOITIVN-PLRYSAKETNV-NISGNGDEGTTIIDSTIIKRYKVGONLNPDSNRITDY 441
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 357 RNMITYVNPQKTYKRETFVNTLT-----GYKFPNDAKFKIYEVTDOQFVDS-FTPD 410
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 442 SEVEDYTN-DYAOLGNNNDVINFGN----IDSPYIIRKIVISKYDPNKDDYTTIQOTYTM 496
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 411 SKLKDVTFQFDVYISNDNKTATVLDLNGOSSDKQYIIIOVAVYPPNNSSTDNKIDYTTLET 470
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 497 OTTINEYGEFTFASVDNTIAFSTSSGQGGDLPEPKTKIGLYWEDVDKDGIONTNDN 556
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 471 QNGKSSWN-----SYSNVNGSSTANGD-----QKKYNLGDYVWEDTNKKGKQDA--N 516
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 557 EKPLSNVLTLTYPDGTS-KSVRTDEDEKYPQFDGV 590
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 517 EKIGKGVYIILKDSNGKELDRTTTDBNGKYQFTGL 551
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

```

RESULT 7
 S41539
 fibrinogen-binding protein - Staphylococcus aureus
 N:Alternate names: clumping factor
 C:Species: Staphylococcus aureus
 C:date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 15-Oct-1999
 C:Accession: S41539; S36630
 R:McDevitt, D.; Francois, P.; Vaudeaux, P.; Foster, T.J.
 Mol. Microbiol. 11, 237-248, 1994
 Title: Molecular characterization of the clumping factor (fibrinogen receptor) of Stap

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 30, 2002, 10:28:50 ; Search time 85.16 Seconds
(without alignments)
773.447 Million cell updates/sec

Title: 1
us-09-147-405-11

Perfect score: 3087
Sequence: 1 HHHHHHPSDEKNDVNNN.....SKSVTDEDGKTPDGVVD 593

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: A_Geneseq_032802.*
2: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1980.DAT:*
3: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1981.DAT:*
4: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1982.DAT:*
5: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1983.DAT:*
6: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1984.DAT:*
7: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1985.DAT:*
8: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1986.DAT:*
9: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1987.DAT:*
10: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1988.DAT:*
11: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1989.DAT:*
12: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1990.DAT:*
13: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1991.DAT:*
14: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1992.DAT:*
15: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1993.DAT:*
16: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1994.DAT:*
17: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1995.DAT:*
18: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1996.DAT:*
19: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1997.DAT:*
20: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1998.DAT:*
21: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1999.DAT:*
22: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA2000.DAT:*
23: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA2001.DAT:*
24: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA2002.DAT:*
25: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA2003.DAT:*
26: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA2004.DAT:*
27: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA2005.DAT:*
28: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA2006.DAT:*
29: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA2007.DAT:*
30: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA2008.DAT:*
31: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA2009.DAT:*
32: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA2010.DAT:*
33: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA2011.DAT:*
34: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA2012.DAT:*
35: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA2013.DAT:*
36: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA2014.DAT:*
37: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA2015.DAT:*
38: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA2016.DAT:*
39: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA2017.DAT:*
40: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA2018.DAT:*
41: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA2019.DAT:*
42: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA2020.DAT:*
43: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA2021.DAT:*
44: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA2022.DAT:*
45: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA2023.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3015	97.7	1092	AAW41602	Staphylococcus epi
2	2812.5	91.1	991	AAW81171	Cell wall protein
3	2812.5	91.1	991	AAW70120	Staph. epidermidis
4	1146.5	37.1	1166	AAW08643	S. aureus SdrE pro
5	857	27.8	22	AAW82803	S. epidermidis ope
6	576	18.7	1315	AAW08642	S. aureus SdrD pro
7	576	18.7	1349	AAW34402	Staphylococcus aur
8	576	18.7	1349	AAW37544	Staphylococcus aur
9	506	16.4	1802	AAW81170	Cell wall protein
10	506	16.4	1802	AAW70119	Staph. epidermidis
11	504	16.3	1155	AAW82343	S. epidermidis ope

12	500.5	16.2	932	22	AAW34082	Staphylococcus aur
13	500.5	16.2	932	22	AAW36845	Staphylococcus aur
14	456	14.8	841	22	AAW34283	Staphylococcus aur
15	456	14.8	841	22	AAW37158	Staphylococcus aur
16	456	14.8	930	20	AAW08641	S. aureus SdrC pro
17	454	14.7	936	18	AAW89801	Staphylococcus aur
18	452	14.6	933	21	AAW58435	Staphylococcus aur
19	452	14.6	933	22	AAW69508	Staphylococcus aur
20	449	14.5	1021	22	AAW33975	Staphylococcus aur
21	449	14.5	1021	22	AAW36951	Staphylococcus aur
22	401.5	13.0	767	22	AAW08640	S. aureus ClfB pro
23	376	12.2	767	22	AAW34403	Staphylococcus aur
24	376	12.2	767	22	AAW37547	Staphylococcus aur
25	373	12.1	940	11	AAW07070	Fibrinogen-binding
26	359.5	11.6	238	18	AAW28019	Staphylococcus aur
27	349	11.3	345	19	AAW31555	Fibrinogen-binding
28	319.5	10.3	978	22	AAW33960	Staphylococcus aur
29	319.5	10.3	1001	22	AAW37093	Staphylococcus aur
30	309.5	10.0	1018	22	AAW34301	Staphylococcus aur
31	309.5	10.0	1018	22	AAW37245	Staphylococcus aur
32	304	9.8	1027	18	AAW89806	Staphylococcus aur
33	301	9.8	1018	9	AAW82115	Fibrinogen binding
34	259.5	8.4	1112	20	AAW08603	S. pyogenes SFBP-
35	219	7.1	537	7	AAW60452	Sequence of the As
36	214	6.9	1279	22	AAW83047	S. epidermidis ope
37	213.5	6.9	2206	21	AAW18254	Plasmodium falcipa
38	208.5	6.8	251	18	AAW89804	Staphylococcus aur
39	206.5	6.7	2368	22	AAW34139	Staphylococcus aur
40	206.5	6.7	2368	22	AAW36796	Staphylococcus aur
41	204.5	6.6	2727	22	AAW61674	Drosophila melanog
42	204	6.6	5024	22	AAW82935	S. epidermidis ope
43	202.5	6.6	1997	21	AAW18287	Plasmodium falcipa
44	202	6.5	1308	21	AAW18167	Plasmodium falcipa
45	201	6.5	665	21	AAW18278	Plasmodium falcipa

ALIGNMENTS

RESULT	ID	Description	Score	Query Match	Length	ID	Description
1	AAW41602	Standard; Protein; 1092 AA.	3015	97.7	1092	AAW41602	Standard; Protein; 1092 AA.
2	AAW41602	22-JUN-1998 (first entry)	2812.5	91.1	991	AAW81171	Cell wall protein
3	AAW41602	22-JUN-1998 (first entry)	2812.5	91.1	991	AAW70120	Staph. epidermidis
4	AAW41602	22-JUN-1998 (first entry)	1146.5	37.1	1166	AAW08643	S. aureus SdrE pro
5	AAW41602	22-JUN-1998 (first entry)	857	27.8	22	AAW82803	S. epidermidis ope
6	AAW41602	22-JUN-1998 (first entry)	576	18.7	1315	AAW08642	S. aureus SdrD pro
7	AAW41602	22-JUN-1998 (first entry)	576	18.7	1349	AAW34402	Staphylococcus aur
8	AAW41602	22-JUN-1998 (first entry)	576	18.7	1349	AAW37544	Staphylococcus aur
9	AAW41602	22-JUN-1998 (first entry)	506	16.4	1802	AAW81170	Cell wall protein
10	AAW41602	22-JUN-1998 (first entry)	506	16.4	1802	AAW70119	Staph. epidermidis
11	AAW41602	22-JUN-1998 (first entry)	504	16.3	1155	AAW82343	S. epidermidis ope

QY 308 FTQYVDKYEIKAHKLTISYIDSKYVNNNTKIDVEYKTAALSSVNTTIYVEORENMENT 367
DB 405 ftyvkvkyenikahkltisylidskypnnntkldveyktaalsvntltiveyqkpnent 464
QY 368 ANQSMFTNIDTAKNHTVEQITINPLRYSAKETNVAISNGDGSSTIIDSTIIRKYKYG 427
DB 465 anqsmftndctknhtveqitlynplrysakentvnisngdgsstliddstlikkykyg 524
QY 428 DNGNLPDSNRNIYDSEYEDVTNDYQALGNNNDVINFGNIDSPYIKYISKYDPKRDY 487
DB 525 dngnlpdsnriydyseyedvtnddyaglgnnndvinfgnidspylkkyiskydpnkddy 584
QY 488 TTIIDQVVTMOTINEXYTGERTASNDYNTAFSSSGOGGDDAPERTYKIGDYVWEDVK 547
DB 585 ttiidqvvtmotinexygtetrassyndntafstssgqgddpptyklygdyvwedvdk 644
QY 548 DGIQNTNDNEKPLSNVLTYPDGTSGSVRTDEDKYOFDGVQ 591
DB 645 dgiqntndnekplsnvltlypdgtsgsvrtdeegkyqfdgik 688

SUPT 4
AY08643
ID AAY08643 standard; Protein; 1166 AA.
XX AAY08643;
AC
XX
DT 09-AUG-1999 (first entry)
XX
DE S. aureus Sdrc protein.
XX
XX Fibrinogen-binding protein: alpha chain; beta chain; ClfB; Sdrc; Sdrc;
KM Sdrc; fibrinogen; medical device; competitive inhibitor; pharmaceutical;
KM treatment; infection; septicemia; osteomyelitis; mastitis; endocarditis;
KM extracellular matrix; vascular graft; vascular stent; vaccine;
KM intravenous catheter; artificial heart valve; cardiac assist device;
antibacterial.
XX
OS Staphylococcus aureus.
XX
XX WO927109-A2.
XX
XX 03-JUN-1999.
XX
XX 25-NOV-1998; 98WO-US25246.
XX
XX 31-AUG-1998; 98US-0098427.
XX
XX 26-NOV-1997; 97US-0066815.
XX
XX (EIDH/) EIDHIN D. N.
XX (FOF-) FORPAS T/A BIORESEARCH IRELAND.
XX (FOST/) FOSTER T. J.
XX (HOOK/) HOOK M. A. O.
XX (INH-) INHIBITEX INC.
XX (JOSE/) JOSEFSSON E.
XX (PAT/) PATRI J. M.
XX (PERK/) PERKINS S. E.
XX
PI Eldhin DN, Foster TJ, Hook MAO, Josefsson E, Patti JM;
PI Perkins SE;
DR MPI: 1999-357844/30.
DR N-PSDB: AAX77594.
XX
XX Staphylococcus aureus fibrinogen-binding proteins for treating
PT septicemia, osteomyelitis, mastitis or endocarditis
XX
XX Claim 8; Fig 9; 143pp; English.
XX
XX This invention describes novel Staphylococcus aureus fibrinogen-binding
CC proteins that bind both the alpha and beta fibrinogen chains. The
CC proteins (and their encoding nucleic acids are ClfB, Sdrc, Sdrc and
CC Sdrc). Staphylococcus aureus is thought to utilize fibrinogen to adhere

CC to medical devices, binding proteins that bind both the alpha and beta
CC fibrinogen chains (ClfB, Sdrc, Sdrc and Sdrc) can therefore be used as
CC competitive inhibitors to block this binding. Antibodies against ClfB,
CC Sdrc, Sdrc and Sdrc inhibit ClfB, Sdrc, Sdrc and Sdrc mediated binding.
CC The proteins of the invention can be used in a pharmaceutical composition
CC for the treatment of Staphylococcus aureus infection e.g. septicemia,
CC osteomyelitis, mastitis or endocarditis or to inhibit the binding of
CC S. aureus to the extracellular matrix. The proteins or their fragments
CC may be used to coat a medical device to reduce the S. aureus infection of
CC an indwelling medical device, especially where the medical device is
CC selected from the group consisting of vascular grafts, vascular stents,
CC intravenous catheters, artificial heart valves, and cardiac assist
CC devices. ClfB, Sdrc, Sdrc, Sdrc, or an active fragment, subdomain or
CC encoding gene may be used as a vaccine. The DS (aspartate serine) repeat
CC region or a gene encoding it may be used as an identifying probe for the
CC identification of genes and encoding proteins from Staphylococcus aureus
CC (other than ClfA), S. hemolyticus, S. lugdunensis, and S. schleierferti.
CC The proteins of the invention have antibacterial activity.
XX
SQ Sequence 1166 AA:
QY 8 SSDEKNDYINNQSINTDNNQI--IKKETNNYDGIKRSERDRETTNDENATPL 65
DB 66 atsdnkevseletensntensnlpkke--tndsqpeakkesstsqkqnnvtat 123
QY 66 OKTPDONTLHEEYKES---SSSESSNSIDAOQPSHTTINREESVQTSQVDESHS 122
DB 124 tekpgn--lkenvpnsdktatetvyllekxapnt--ndvltkps-----ts 172
QY 123 DRANSKIKESNT---SGREENTIEQPNKYKEDSTTSQPSGYNNIDEKIS--MODELNLP 178
DB 173 epstselqkptpgestnlensqppkpskd---nqytdenpknepnvskelkmp 229
QY 179 -----INYEKARPLSTTSQAOPSTKR-----TYNQLAAGQSNVHLIYVTDQ 223
DB 230 ekikelvndnsndhsktpatavapkrvaknrfavaqpaavaasnvdllkvtkx 289
QY 224 SITEGYDSEGYKADAENLIDYVTEFVDDKVKSGDITVDDKNTVPSDLTFTIPK 283
DB 290 tlkvq-dgkxnaaahdgkdietydteftldnkvkksqtdmtnlydknvlpeditdkndpid 348
QY 284 IKDNGEIIATGYDNKNQIITFTDYDYKYEIKAHKLTISYIDSKYVNNNTKLDVE 343
DB 349 ltdpsgevlakgtfdkackqlytftcdydkyediksriltlyyidkkvyp-netalnt 407
QY 344 YKTAALSSVNTTIYVEORENMENTANLOSFTNIDTNRNHTVEQITINPLRYSAKETNVA 403
DB 408 fatagketeqnvltvdyqdmvnhgsnlqstlftldedkqtleaglyvnpklsakutlkvd 467
QY 404 ISNG-----DEGSTIIDSTIIRKYKYVNDONLPLPSNRNIYDSEYEDVTND-DYA 453
DB 468 lagsqdydgnlklngstltdqnteklykvnssdqglpnsnlyfsgyedtsgfdnk 527
QY 454 QLGNNNDVINFGNIDSPYIKYISKYDPKRDYTTIIOQTVMOTINEXYTGERTASTD 513
DB 514 NTIAFTSSGOGGDD-PEPKYKIDDYWEDVDKGIQNTNDNEKPLSNVLTLYPPG 572
QY 586 nftvsnldcgggdyvkvpeklykgdywedvdkdygvtdekekpmavnlvltlypdg 645
QY 573 TSKSVRTDEDKYOFDGVQ 591
DB 646 tksvrtdeaghyefgik 664

RESULT 5
AAG82803

ID AAG82803 standard; Protein: 278 AA.
 AC AAG82803;
 XX
 DT 03-SEP-2001 (first entry)
 XX
 DE S. epidermidis open reading frame protein sequence SEQ ID NO:2700.
 XX
 KM Staphylococcus epidermidis SRI strain; infection: diagnosis;
 KM vaccination; endocarditis.
 OS Staphylococcus epidermidis.
 XX
 PN WO200134809-A2.
 XX
 PD 17-MAY-2001.
 XX
 PF 09-NOV-2000; 2000MO-US10782.
 XX
 V 09-NOV-1999; 99US-0164258.
 PA (GLAX) GLAXO GROUP LTD.
 XX
 FI Kimmberly MJ;
 XX
 DR WPI; 2001-316495/33.
 DR N-PSDB; AAH53653.
 XX
 PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
 PT useful for vaccinating against infections, e.g. endocarditis -
 PS
 PS Claim 18; Page 705; 2188pp; English.
 XX
 XX AAH53970 represent nucleic acids (I) encoding polypeptides
 CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
 CC (I) and (II) can have antibacterial activity and therefore can be used
 CC in vaccination. The nucleic acids (I) may be used to produce the
 CC S. epidermidis polypeptides (II) via the production of vectors
 CC containing them which are used to produce hosts cells which express the
 CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
 CC used to vaccinate subjects and to raise antibodies against the bacteria.
 CC The polypeptides may also be used to assay for other inhibitors of their
 CC activity and therefore identify compounds that may be used for the
 CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
 CC AAH55090 represent specifically claimed S. epidermidis genomic DNA
 CC polynucleotide sequences from the present invention. AAH55091 to
 CC AAH55098 represent oligonucleotide sequences and primers which are used
 CC in the exemplification of the present invention.
 CC N.B. The present invention specifically claims all the polynucleotide
 CC sequences given in the sequence listing of the present specification,
 CC however the sequence listing only goes up to SEQ ID NO:4454 so even
 CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
 CC no sequences are present for SEQ ID NO:4455 to 4464.
 CC
 XX
 XX Sequence 278 AA:
 SQ
 Query Match 27.8%; Score 857; DB 22; Length 278;
 Best Local Similarity 93.0%; Pred. No. 4.4e-39;
 Matches 172; Conservative 3; Mismatches 8; Indels 2; Gaps 1;

QY 188 PLSTT 192
 DB 253 KLSVT 257
 RESULT 6
 ID AAY08642
 AAY08642 standard; Protein: 1315 AA.
 XX
 AC AAY08642;
 XX
 DT 09-AUG-1999 (first entry)
 XX
 DE S. aureus SdrD protein.
 XX
 XX Fibrinogen-binding protein; alpha chain; ClfB; SdrC; SdrD;
 KM SdrE; fibrinogen; medical device; competitive inhibitor; pharmaceutical;
 KM treatment; infection; septicemia; osteomyelitis; mastitis; endocarditis;
 KM extracellular matrix; vascular graft; vascular stent; vaccine;
 KM intravenous catheter; artificial heart valve; cardiac assist device;
 KM antibacterial.
 KM
 XX Staphylococcus aureus.
 OS
 XX WO9927109-A2.
 XX
 PD 03-JUN-1999.
 XX
 PF 25-NOV-1998; 98WO-US25246.
 XX
 PR 31-AUG-1998; 98US-0098427.
 PR 26-NOV-1997; 97US-0066815.
 XX
 PA (EIDH/) EIDHIN D N.
 PA (FOF/) FORFAS T/A BIORESEARCH IRELAND.
 PA (FOST/) FOSTER T J.
 PA (HOOK/) HOOK M A O.
 PA (INH/) INHIBITEX INC.
 PA (JOSE/) JOSEFSSON E.
 PA (PAT/) PATTI J M.
 PA (PERK/) PERKINS S E.
 XX
 PI Eidlun DN, Foster TV, Hook MAO, Josefsson E, Patti JM;
 PI Perkins SE;
 XX
 DR WPI; 1999-357844/30.
 DR N-PSDB; AAX77593.
 XX
 PT Staphylococcus aureus fibrinogen-binding proteins for treating
 PT septicemia, osteomyelitis, mastitis or endocarditis
 PS
 PS Claim 8; Fig 8; 143pp; English.
 XX
 CC This invention describes novel Staphylococcus aureus fibrinogen-binding
 CC proteins that bind both the alpha and beta fibrinogen chains. The
 CC proteins (and their encoding nucleic acids are ClfB, SdrC, SdrD and
 CC SdrE). Staphylococcus aureus is thought to utilize fibrinogen and beta
 CC to medical devices, binding proteins that bind both the alpha and beta
 CC fibrinogen chains (ClfB, SdrC, SdrD and SdrE) can therefore be used as
 CC competitive inhibitors to block this binding. Antibodies against ClfB,
 CC SdrC, SdrD and SdrE inhibit ClfB, SdrC, SdrD and SdrE mediated binding.
 CC The proteins of the invention can be used in a pharmaceutical composition
 CC for the treatment of Staphylococcus aureus infection e.g. septicemia,
 CC osteomyelitis, mastitis or endocarditis or to inhibit the binding of
 CC S. aureus to the extracellular matrix. The proteins or their fragments
 CC may be used to coat a medical device to reduce the S. aureus infection of
 CC an indwelling medical device, especially where the medical device is
 CC selected from the group consisting of vascular grafts, vascular stents,
 CC intravenous catheters, artificial heart valves, and cardiac assist
 CC devices. ClfB, SdrC, SdrD, SdrE, or an active fragment, subdomain or
 CC encoding gene may be used as a vaccine. The DS (aspartate serine) repeat
 CC region or a gene encoding it may be used as an identifying probe for the
 CC identification of genes and encoding proteins from Staphylococcus aureus

CC (other than C1A), *S. hemolyticus*, *S. lugdenensis*, and *S. schleriferi*
CC The proteins of the invention have antibacterial activity.
xx

Query Match	18.7%;	Score 576;	DB 20;	Length 1315;
Best Local Similarity	30.1%;	Pred. NO. 5.3e-23;		
Matches 186;	Conservative 112;	Mismatches 228;	Indels 92;	Gaps 29

QY	8	SSDEKRDV---INNNOISINTDNNQILKKEETNNVGIKRSRDRRESTINDE-NEAT	63
Db	55	stdehleaatsadsdvdmgqingetntkndqkemsvsgqntsgnkllheke	114
QY	64	FLQKTPDNDTHLEEEEEKSESSVESSNSSIDTAQAQPSHTIINRESVQTSQDNVDSHVD	123
Db	115	svqstignkvsvstaksdeqaspskstmedntkq-----clsngeaig-pdlgenksvvn	168
QY	124	FANSKIEESTESCKEENTIEQPNKVKEDSTTSQPSGYNI-DEKISMODEL--NLPIN	160
Db	169	-----vgptneenkvd-----aktest-----lnvkedaalkendetlvdmns	209
QY	181	EYENKARPLSTTSQAP-----SIKRVYVNOQLAADGQSVNHLIKVTDOSTIEGIDDSGCVI	236
Db	210	nnennadilipksetaprlntmrtaavgssteaknvnlltsntltlvdadkmkiv	269
QY	237	KAHDAENLIDVFEFVEDDKVKSQDTMTVDIDKNTV-----PSDLTPSFTIPKID-NSG	289
Db	270	pedqylslksgqit--vdckxksqdyfiky-sdvqvyyglnpedlkn---lqdklpnng	323
QY	290	EIIATGTYNKKNQIOTYTFEDVDYDKYENIKAHKLTSYIDKSKVPNNNTKLADVEYKTALS	349
Db	324	etlatahdcannllytfcdydrfnsvgmqlnyslymdactlp--vskndvefnvllg	381
QY	350	SVAKTIVEYQRPN--ENFRANLQSMPTNIDTKNHT-----VEDTIIINFLRSAKE	399
Db	382	ntctcttanlqydyvvnneknsigsatf--elvshgnkenpgykqilyvnpnsnltm	439
QY	400	TNNVI-----SGNDEGSTIIDSTIIIKVYKVGONQNLPSNRIYDYS--EYEDVTNDQVA	453
Db	440	aklkvqgahysygnmngqlnkvtldkllyvprgylt---nkgydvntkeltdvtn-qyl	485
QY	454	Q---LGNNDVNIINFGNIDSPYIIKVIKSYDPKKDDYTTIQGTVMQTINNEYTGEFRTA	510
Db	496	qkltgndmsaavidfgnadasyvwmvntkfygnsepsltlvmaltlssign-----ksv	549
QY	511	SYDNTIAFSNSSGQGGDLPPEKTYTKIGDVMVEDVDKDGQINNDNEKPSANLVLTITP	570
Db	550	stgnalqfctngsgggg---gevykignywedtnknvgel--gekgyvntlevtvd	602
QY	571	DGTSKSYR---TDDEGKY	585
Db	603	nnctktygeavtkedgisy	620
RESULT 7			
AAU34402	ID	AAU34402 standard; Protein; 1349 AA.	
AAU34402:			
14-FEB-2002		(first entry)	
Staphylococcus aureus		cellular proliferation protein #678.	
Antisense; prokaryotic cellular proliferation protein;			
antibiotic; antibacterial; drug design.			
Staphylococcus aureus.			
MO2001070955-A2.			
27-SEP-2001.			

```

XX XX 21-MAR-2001; 2001IWO-USO9180.
PR PR 21-MAR-2000; 2000US-191078P.
PR PR 23-MAY-2000; 2000US-206848P.
PR PR 26-MAY-2000; 2000US-207727P.
PR PR 23-OCT-2000; 2000US-242578P.
PR PR 27-NOV-2000; 2000US-253625P.
PR PR 22-DEC-2000; 2000US-257931P.
PR PR 16-FEB-2001; 2001US-269308P.
PA PA (ELIT-) ELITRA PHARM INC.
XX XX
PI PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Traxwick JD, Carr GT;
PI Yamamoto RT, Xu HH;
DR DR WPI, 2001-611495/70.
DR N-PSDB; AAS52261.
PT PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX XX
XX XX Example 3; Seq ID No 5898; 511pp; English.
CC CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins'
CC and to obtain antibodies capable of binding to the expressed proteins'.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX XX
SQ Sequence 1349 AA;
Query Match 18.7%; Score 576; DB 22; Length 1349;
Best Local Similarity 30.1%; Pred. No. 5.5e-23;
Matches 186; Conservative 112; Mismatches 228; Indels 92; Gaps 29;
QY 8 SSDSEKNDV--INNOSINTDDNNQIKKEETNYDGIKESSEDRTSTNWDE--NEAT 63
DY ::::|:|:-:||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
DB 55 stnkelneatstsaadngsdskvdmqnglndektnhdnkmgwssgnettsngkhlieke 114
QY 64 FLAKTPDNTHLTREEVKESSVESSNSIDTAQPSPHTTINREESVOGTSNVEDSHVSD 123
DY :|:~|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
DB 115 svsgltgnkvkstakdsdeqaspksntedlnltkq-----tlisngalg-pdigenkxsvn 168
QY 124 FANSKIESNTSESKKEENTIEQPKNVKEDSTTSQPSGTNI-DKETSODELL--NLPIN 180
DY :~::~:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
DB 169 -----vprhtneenkvd-----akeestl-----lrvksdaiknsdelivdnans 209
QY 181 EYENKAPRLSTSAQP-----SIKRVTVAQLAEAGSNVNHILKVDAQSITREGYDSEGVY 236
DY ||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
DB 210 mnenadiillpkstaprlntlmrlaavqpssteakenvnditsttlltvaddaknnkiv 269
QY 237 KAHDAEWLITDVFEVDKRYKSGETMIVYDIKNTV-----PSDIFDSFTIPRKID-NSG 289
DY :||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
DB 270 padgdylslksqilt--vdckvksgdyftlkj-sdctvgvglnpedi kn--lgdlkdpng 323
QY 290 EIATGATGYDKKKNGQITTFPDYVQKYENIKAHKLFTSYIDSKVYNNMTXIDVYEKTALS 349

```

```

Query Match      18.7%; Score 576; Db 22; Length 1349;
Best Local Similarity 30.1%; Pval. No. 5_e-23;
Matches 166; Conservative 112; Mismatches 228; Indels 92; Gaps

QY      8  SSDEEKDV---INNOSINTDDNNQIIEETNWDGIERSEDRTESTYNWE-NEAT 63
       |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      55 stklneatssadngsdskvdmqgingednctkndkgkewavsgneltsgnkllheke 114
       |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY      64 FLTKPDODNTHLREEEKESSVESSNSSIDTQAQPESHITINNEESVQTSDNNEDSVSD 123
       |:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db     115 svastgcnkvkstakdsdegaaspstnedintky----tlisqealq-pdlgenksavn 168
       |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY     124 FANSKIKESNTSEGKEENTTEOPNKVKVEDSTTSPPSGTYNI-DEKISMODEL--NLPIN 180
       |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db     169 -----vgpltnenkdkvd-----akteest-----lnvsdaalksnleelvdnnms 209
       |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY     181 EYENKARPLSTSNAP-----SIKRYYVYNQLAAEGSVMNHLIKVTDSITEGYDSEGVY 236
       |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db     210 mnmehedlilpkstaprlntlmrtlaavgpssteakvnndllfsntltlvdadkmnkiv 269
       |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY     237 KAHDAENLIDVFEEVDQDKVSGDTMTVIDDKMNV-----PSDLTFSPFPIKD--NSG 289
       |:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db     270 pcdgyalsysqlt--vdckkxsgdyfilyk-sdtvyvygjnepedikh---ldgatkpnmj 323
       |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY     290 ETIATGTGNKKKOIIYTPFDYDYDKENIKAHLKLTSYIDKSVPNNNTKLIVGYTTALS 349
       |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

```

Cc 324 elatakhdaannllyftdydrfnsvqmglnyslymadclp--vskndvefnvlig 381
Cc 350 SVNKTIIVEYQRPN--ENRTANLOSMTNIDTKNHT-----VEQTIYINPLRYSAKE 399
Cc 382 ntctktanlgydpdyvneknsgisaf--etvshvgnkenpgyykqtllynpensaln 439
Cc 400 TNVNI-----SGNGDEGSTIIDSTIIKVKYKGNQNLPSNRRIYDYS--EYEDVTNDYA 453
Cc 440 akikvqayhsypnlglnkdvcltklygvpkgytl---nkgydvntkcltdvtn-qy1 495
Cc 454 Q---LGNNDVNIINGNIDSPYIIKVKYKGNQNLPSNRRIYDYS--EYEDVTNDYA 510
Cc 496 qkltvgdnnsavldfgnadasyvwmvntkfytnsepspllvqmatlsstgn-----ksv 549
Cc 511 SYDNTIAFSTSSGOGGDLPEPKTYKIGDYVEDVDKDIQNTNDNEKPLSNVLTLYTP 570
Cc 550 stgnalgtfnngsgag-----gevyklygnvwdtnknvgel--gekvgvntvltv-fd 602
Cc 571 DGTSSKSVR---TDEDGY 585
Cc 603 nntnkvgeavtkedggy 620

RESULT 8
AAU37544 standard; Protein; 1349 AA.
XX
AC AAU37544;
XX
Dt 14-FEB-2002. (first entry)
XX
DE Staphylococcus aureus cellular proliferation protein #1714.
XX
KW Antisense: prokaryotic cellular proliferation protein;
KW antibiotic; antibacterial; drug design.
XX
OS Staphylococcus aureus.
XX
PN W0200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 23-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DR WPI: 2001-611495/70.
XX
N-PSDB: AAS55403.
XX
XX New polynucleotides for the identification and development of
XX antibiotics, comprise sequences of antisense nucleic acids .
XX
XX Example 3; Seq ID No 13137; 511pp; English.
XX
Cc The invention relates to antisense inhibitors of genes essential to
Cc prokaryotic cellular proliferation, their use in identifying the
Cc genes, their use in the discovery of novel antibiotics, the essential
Cc genes themselves and the encoded proteins. The prokaryotes used are
Cc Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
Cc pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
Cc invention is also useful for the identification of potential new targets
Cc for antibiotic development. The antisense nucleic acids can also be used

```

to identify proteins used in proliferation, to express these proteins, and to obtain antibodies capable of binding to the expressed proteins. The proteins can be used to screen compounds in rational drug discovery programmes. The antisense nucleic acid sequence is also useful to screen for homologous nucleic acids which are required for cell proliferation in a wide variety of organisms. The present sequence represents an essential prokaryotic cellular proliferation protein.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).

Sequence 1349 AA:

Query Match 18.7%; Score 576; DB 22; Length 1349;
 Best Local Similarity 30.1%; Pred. No. 5,56-23;
 Matches 186; Conservative 112; Mismatches 228; Indels 92; Gaps 29;

```

Cc 8 SDEEKNDV---INNOSINTDNNQIIRKEETNNYDQIEKRSDEKSTTNDE-NEAT 63
Cc 55 stnkelnatetsaadsngsdvmdqngdnkndkngkemsyqgneltngnklleke 114
Cc 64 FLOKTPDQNTHLTEEEVSSSVSSNSIPTAQQPSHTTINREESVQTSNVEDSHYSD 123
Cc 115 svsgltgnkvevstaksdeqaspxstnedlntkg-----tlngaalq-palgenksvyn 168
Cc 124 FANSKIKESNTESGKEENTIEQPNKVKEDSTTSQPSGYTN-DEKISQDELL--NLPIN 180
Cc 169 -----vqplneenkkyd-----akestt-----lnvksdakandellvdnms 209
Cc 181 EYENKARPLSTTSQAP----SIKRYTVNQLAEDOSNNVHLIKYTDQSITEGYDSEVI 236
Cc 210 nneennadllpkstapklntmrtaavgssteeknvndlltntcltvvadknkiv 269
Cc 237 KAHDAENLIYDTEFVDKVKSGDTMTVDIDKNFY-----PSDLTDSFTTPKIKD-NSG 289
Cc 270 pagdylslksqit--vddkvnsgdyftlky--sdvqvyglnpedkn---lgdldpnnq 323
Cc 324 elatakhdaannllyftdydrfnsvqmglnyslymadclp--vskndvefnvlig 381
Cc 350 SVNKTIIVEYQRPN--ENRTANLOSMTNIDTKNHT-----VEQTIYINPLRYSAKE 399
Cc 382 ntctktanlgydpdyvneknsgisaf--etvshvgnkenpgyykqtllynpensaln 439
Cc 400 TNVNI-----SGNGDEGSTIIDSTIIKVKYKGNQNLPSNRRIYDYS--EYEDVTNDYA 453
Cc 440 akikvqayhsypnlglnkdvcltklygvpkgytl---nkgydvntkcltdvtn-qy1 495
Cc 454 Q---LGNNDVNIINGNIDSPYIIKVKYKGNQNLPSNRRIYDYS--EYEDVTNDYA 510
Cc 496 qkltvgdnnsavldfgnadasyvwmvntkfytnsepspllvqmatlsstgn-----ksv 549
Cc 511 SYDNTIAFSTSSGOGGDLPEPKTYKIGDYVEDVDKDIQNTNDNEKPLSNVLTLYTP 570
Cc 550 stgnalgtfnngsgag-----gevyklygnvwdtnknvgel--gekvgvntvltv-fd 602
Cc 571 DGTSSKSVR---TDEDGY 585
Cc 603 nntnkvgeavtkedggy 620

RESULT 9
AAV83170
ID AAV83170 standard; Protein; 1802 AA.
XX
AC AAV83170;
XX
Dt 24-JUL-2000 (first entry)
XX
DE Cell wall protein SdrF.

```

Dd	291	tklkpvsrtdssvndk--qgyitrsav-----asjgvsdseneatlnavrhdldlkaastre	344	aa
Qy	163	NIDEKI-----SNODELLNLP--INEXENKARPLSTTSAPQSIKRYTVNQLAAE--	209	aa
Dd	345	qineellaaeklkdfsnpgdyvdcplalnrgsqnsp--hkasaap---rmmlmaiaepn	399	aa
Qy	210	OGSNVNHILKVTDO--SITEGYDSEGEYKAKHDAENLLYDTEFVDDQKVASGDTMYDID	267	aa
Dd	400	sgknvndkvkltnpctlsnksnnhannvlpwtneqfqlkanyelddasikegdtflkyg	455	aa
Qy	268	KNTVPSDLTDSFTIPIKINDNSGEIATGTYDNKKKQITVYTFDVEDVDEYENIKAKHLKTSY	327	aa
Dd	460	qyirpvgjelpaliktqlskdgsilvangvykntcttlyftfnvydqnlqtsfdliat	519	aa
Qy	328	IDKSVPRNNNTKRLDVEYKFTALSSVNNKTIITVEYQRPNEFRNANLQSMFTNIDKNHVEQT	387	aa
Dd	520	pkretalikhnpqymevlianeevkkdfivdygnkkhntlt---aavandvnnkhnev	576	aa
Qy	388	IYINPLRSAKETNVNINISGNGDESTIIDDSTI---IKYKVGQDNQLPQS--NRIVYSE	443	aa
Dd	577	vyllngnnpkyay-----fstvkngelflbgewkkyvevctnamvdsfnpdlssn	628	aa
Qy	444	YEDVTNDNDYAL--GNMNDVNINFGNI---DSPYIIKVIYSKYPKKDDYTTIQQVTMQTT	499	aa
Dd	629	vkdvtsqfaprvsadvsgtrvdlnfarmsangkkylvtgavrvptgtnvyt--eywltrdgt	686	aa
Qy	500	INEXTGEPRRTSYONTIAFSTSSQGGGDLPREKTYKIGDYVWDVDDKDIQNTNDNEKP	558	aa
Dd	687	tn--tnddygskctvrcylvlmgssstaagvdm--tyslgydvldknkngvq--dddekq	739	aa
Qy	560	LSNVLVTLTYTPDQTSKS---VRIDEDGRYQFDGVO	591	aa
Dd	740	lagvyvcl--kdsnmrelgrvttdqsgshygdndq	772	aa
RESULT 10				
ID	AA70119	AA70119 standard; Protein; 1802 AA.		
XX	AA70119;			
XX	06-JUN-2000	(first entry)		
DE	Staph. epidermidis	serine-aspartate repeat region protein SdrF.		
XX	Multiple component vaccine: immunostimulatory; antihapserial; MSCRAM;			
KW	Microbial surface components recognising adhesive matrix molecules;			
KW	collagen binding protein; CBP; CNA; fibrinogen binding protein;			
KW	Clumping factor A; ClfA; Clumping factor B; ClfB; FnBP;			
KW	fibrinectin binding protein; Staphylococcus infection;			
XX	serine-aspartate repeat region protein; SDR protein; SdrF.			
OS	Staphylococcus epidermidis.			
FH	Key	Location/Qualifiers		
FT	Misc-difference 12	/note="Encoded by in-frame stop codon TAA"		
FT	Misc-difference 28	/note="Encoded by in-frame stop codon TAA"		
FT	Misc-difference 1771	/note="Encoded by in-frame stop codon TAA"		
FT	Misc-difference 1771	/note="Encoded by in-frame stop codon TAA"		
FT	Misc-difference 1774	/note="Encoded by in-frame stop codon TGA"		
FT	Misc-difference 1793	/note="Encoded by in-frame stop codon TAA"		
FT	Misc-difference 1800	/note="Encoded by in-frame stop codon TAA"		
XX	W0200012131-A1.	/note="Encoded by in-frame stop codon TGA"		
XX	09-MAR-2000.			
XX	31-AUG-1999;	99WO-US19727.		

```

XX 31-AUG-1998; 98US-0098439.
PR (INH-) INHIBITEX INC.
PA (TEXA-) UNIV TEXAS A & M SYSTEM.
PA (QUEB-) QUEEN ELIZABETH COLLEGE DUBLIN.
XX
PI Patti JM, Foster TJ, Hook M;
DR WPI: 2000-237781/20.
DR N-PSDB; AA251201.
XX
PT Composition used for generating immune response or for inhibiting
PT microbial colonization in an animal comprises antibodies that bind
PT collagen binding protein, fibrinogen binding protein and, optionally,
PT fibronectin binding protein.
XX
PS Claim 8; Fig 3; 115pp; English.
XX
The patent discloses multicomponent vaccines containing selected
combinations of bacterial binding proteins termed MSCRAM (microbial
surface components recognising adhesive matrix molecules) or their
antibodies. A vaccine composition is provided that includes collagen
binding protein or peptide, e.g. CNA, a fibrinogen binding protein
preferably clumping factor A (ClfA) or clumping factor B (ClfB),
and optionally a fibronectin binding protein e.g. FnbA-A.
The vaccines are useful for imparting protection against a broad
spectrum of Staphylococcal strains and for inhibiting microbial
colonisation, especially of Staphylococcus aureus, in an animal.
The combinations can also be used to select donor blood pools for the
preparation of purified blood products for passive immunisation.
The present sequence is a serine-aspartate repeat region
protein, Sdr from Staphylococcus epidermidis. The Sdr protein is
useful in vaccine preparation in combination with specific
bacterial binding proteins. These vaccines can be used to treat a broad
spectrum of bacterial infections, including those arising from both
coagulase-positive and coagulase-negative bacteria.
XX
Sequence 1802 AA;
SQ
Query Match 16.4%; Score 506; DB 21; Length 1802;
Best Local Similarity 27.4%; Pred No. 4.9e-19;
Matches 174; Conservative 101; Mismatches 276; Indels 84; Gaps 27;
XX
7 PSSDEE--KNDVNNNSINTDNNQIIRKEETNYDGIKRSDDRTST----- 54
172 pvneseiaetpkrstetqdsteknpsl-kdnlnassstskedstckqagmstnk 230
55 TIVDENEA-TFLOKT-PQONTHTTEEVKSSSVSSNS-----SIDTAQOQSHTT 103
231 snldtdpctqsktsqandstndqspksqldskpsqkytkkfndepdqvtelt 290
104 IN-REESVOTSDNVEDSHSDFANSKRKIKSNTESGKEENTIEQPHVKVEDSTTSQPSGYT 162
291 tklktpsvstsdsvnk--qdytrsav-----aslygdsnetetlnavrdnlhkaastre 344
163 NIDEKT-----SNOBELMLP--INIEYENKARPLSTTSQPSIKRYTNOQLAAE-- 209
345 qthleaaalkkdfanpdyvtdplalnsgksnsp--hksasp--rmlnmslaaepn 399
210 OGSNVAHLIKVMDQ--SIEGVDSDSGVJKAHAENLIYVFEVDKVSQSGDTMYVID 267
400 sgfnvndkxklmptclslmsnhanhvwptsnegfnlkanyleddslkgdftlkgy 459
268 KNTVPSDLTDSFTPIKIDNSGEIATGTYDNKNKOITTYFTYDYKYENIKALKLTSY 327
460 qylrpgjlelpaikqlrksdgsivangydknttityftnyvdyqqltgsfdlat 519
328 IDKSKYPNNNTKIDVEKTAALSSVNTTIVEYORPNENRANTIQSMFTNIDTKNHTVEQT 387
520 pkretaikdnqymevrtlanevkkdfivdygnkkdntt---aaavanvnnknknev 576

```

```

QY 388 IYINPLRYSAKETNWNISGNGEGSTIIDSTI---IKYKVGDNQNLPSD-NRIYDSE 443
Db 577 vylngnngpkkyak-----fstvkngelfipevkvewtdcnamdsfnpalnsn 628
QY 444 YEDVTNDDYAQL-GNNNDVYNFGNI--DSPYIKVISKYDPNKKDYTTIOQVYTMQTT 499
Db 629 vkdvtsqfapksadgltrdlnfarsmangkkyivlgavrpqfignvle--eywltrdgt 686
QY 500 INEYGEFFRTASDNTIAFSTSSGQGGDLPEPKYTKIDYIWDYDKGICIONTDNEXP 559
Db 687 tn--tndfyrgtkstvtlylngssstaggdp--tysldywdlknkngv--dddexg 739
QY 560 LSNVLVTLYTPDGTSSK---VRTDEDKXQFQGVQ 591
Db 740 lagvyvtl--kdsnrelgrvtcdgshyqtdnlg 772
XX
RESULT 11
AA682343
ID AA682343 standard; Protein; 1155 AA.
XX
AA682343;
XX
03-SEP-2001 (first entry)
XX
DE S. epidermidis open reading frame protein sequence SEQ ID NO:1780.
XX
Staphylococcus epidermidis SRI strain; infection; diagnosis;
KM vaccination; endocarditis.
XX
Staphylococcus epidermidis.
XX
WO200134809-A2.
XX
17-MAY-2001.
XX
09-NOV-2000; 2000MO-US30782.
XX
09-NOV-1999; 99US-0164258.
XX
(GLAX) GLAXO GROUP LTD.
XX
Kimmerly WJ;
XX
WPI: 2001-316495/33.
XX
N-PSDB; AAH53193.
XX
Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
useful for vaccinating against infections, e.g. endocarditis -
Claim 18; Page 491-492; 2188pp; English.
XX
AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
(II), given in AA681454 to AA683120, from Staphylococcus epidermidis.
CC (I) and (II) can have antibacterial activity and therefore can be used
in vaccination. The nucleic acids (I) may be used to produce the
S. epidermidis polypeptides (II) via the production of vectors
containing them which are used to produce host cells which express the
polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
used to vaccinate subjects and to raise antibodies against the bacteria.
CC The polypeptides may also be used to assay for other inhibitors of their
activity and therefore identify compounds that may be used for the
treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
AAH55090 represent specifically claimed S. epidermidis genomic DNA
polynucleotide sequences from the present invention. AAH5091 to
AAH55098 represent oligonucleotide sequences and primers which are used
in the exemplification of the present invention.
CC N.B. The present invention specifically claims all the polynucleotide
sequences given in the sequence listing of the present specification,
CC however the sequence listing only goes up to SEQ ID NO:4454 so even
CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
no sequences are present for SEQ ID NO:4455 to 4464.
XX

```


50 **Sequence** **1155 AA;**

Query Match	16.38; Score 504; DB 22; Length 1155;
-------------	---------------------------------------

Best Local Similarity 27.28; Pred. No. 3.5e-19,
Matches 173; Conservative 103; Mismatches 376

Series	1/3, Conservative	10/2, Mismatches	2/0, Inlets	84, Gaps	2/1
Series	1/3, Conservative	10/2, Mismatches	2/0, Inlets	84, Gaps	2/1

QY	7	PSSEQ---	KNDV	INNOS	INTD	NNNO	II	KKKEE	NNN	DG	IE	KRE	RED	TE	S	-----	54									
		135	ptvneesa	ae	pk	st	cg	qg	d	ae	kn	psl	-	k	dn	ns	st	ts	ck	ae	q	g	st	n	k	
Db		135	ptvneesa	ae	pk	st	cg	qg	d	ae	kn	psl	-	k	dn	ns	st	ts	ck	ae	q	g	st	n	k	
QY		55	TW	V	EN	A	-	T	F	L	O	K	T	-	P	O	D	N	T	H	L	E	E	V	K	S
		194	snldtnd	s	pt	g	s	e	k	t	s	q	a	m	d	s	r	c	d	n	s	t	d	n	s	t
Db		194	snldtnd	s	pt	g	s	e	k	t	s	q	a	m	d	s	r	c	d	n	s	t	d	n	s	t
QY		104	I	N	-	R	E	E	S	V	O	R	S	D	N	E	D	S	H	V	S	F	A	N	S	K
		254	tkltp	s	t	s	t	s	v	n	d	k	-	q	d	y	t	r	s	a	v	-	a	s	i	g
Db		254	tkltp	s	t	s	t	s	v	n	d	k	-	q	d	y	t	r	s	a	v	-	a	s	i	g
QY		163	N	I	D	E	K	I	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
		308	qineaa	i	la	e	a	k	k	f	s	p	d	y	a	v	d	e	p	l	a	n	c	s	q	a
Db		308	qineaa	i	la	e	a	k	k	f	s	p	d	y	a	v	d	e	p	l	a	n	c	s	q	a
QY		210	Q	G	S	V	N	H	L	I	K	V	T	D	-	S	I	T	E	G	Y	D	S	E	G	V
		363	sgknvnd	k	v	k	l	n	p	l	a	i	n	k	m	a	n	n	v	i	p	t	s	e	g	n
Db		363	sgknvnd	k	v	k	l	n	p	l	a	i	n	k	m	a	n	n	v	i	p	t	s	e	g	n
QY		268	K	N	T	V	P	S	D	L	T	S	E	F	I	P	K	I	K	D	S	G	E	I	A	N
		423	qyitp	g	l	e	p	g	l	e	p	a	k	t	q	i	r	s	k	d	s	i	v	a	n	g
Db		423	qyitp	g	l	e	p	g	l	e	p	a	k	t	q	i	r	s	k	d	s	i	v	a	n	g
QY		328	I	D	K	S	V	E	N	N	N	T	K	I	D	V	E	X	K	T	A	L	S	V	N	K
		483	p	k	r	e	t	a	k	i	d	q	n	g	y	m	e	v	i	a	e	v	k	k	f	i
Db		483	p	k	r	e	t	a	k	i	d	q	n	g	y	m	e	v	i	a	e	v	k	k	f	i
QY		388	I	Y	I	N	P	L	R	S	A	K	E	T	V	N	I	S	G	N	D	E	S	T	I	
		540	v	l	i	n	q	n	p	y	a	k	y	-	-	-	-	-	-	-	-	-	-	-	-	
Db		540	v	l	i	n	q	n	p	y	a	k	y	-	-	-	-	-	-	-	-	-	-	-	-	
QY		444	Y	E	D	V	T	N	D	D	Y	A	O	L	-	G	N	N	D	V	N	I	N	F	G	
		592	v	k	d	r	s	q	f	k	p	s	a	d	g	t	r	v	d	i	n	f	a	r	s	
Db		592	v	k	d	r	s	q	f	k	p	s	a	d	g	t	r	v	d	i	n	f	a	r	s	
QY		500	I	N	E	T	G	E	R	T	A	S	T	O	N	T	A	F	S	S	G	G	G	D	L	
</																										

```

PR      21-MAY-2000; 2000US-191078P.
PR      23-MAY-2000; 2000US-206848P.
PR      26-MAY-2000; 2000US-207721P.
PR      23-OCT-2000; 2000US-242578P.
PR      27-NOV-2000; 2000US-253625P.
PR      22-DEC-2000; 2000US-257931P.
PR      16-FEB-2001; 2001US-269308P.
XX
XX      (ELIT-) ELITRA PHARM INC.
XX
XX      Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Traxick JD, Carr GJ;
PI      Yamamoto RT, Xu HH;
DR      WPI: 2001-611495/70.
XX      N-PSDB; AAS51941.
XX
PT      New polynucleotides for the identification and development of
PT      antibiotics, comprise sequences of antisense nucleic acids -
XX
XX      Example 3; Seq ID No 5578; 511pp; English.
XX
XX      The invention relates to antisense inhibitors of genes essential to
CC      prokaryotic cellular proliferation, their use in identifying the
CC      genes, their use in the discovery of novel antibiotics, the essential
CC      genes themselves and the encoded proteins. The prokaryotes used are
CC      Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC      pneumoniae, pseudomonas aeruginosa and Enterococcus faecalis. The
CC      invention is also useful for the identification of potential new targets
CC      for antibiotic development. The antisense nucleic acids can also be used
CC      to identify proteins used in proliferation, to express these proteins,
CC      and to obtain antibodies capable of binding to the expressed proteins.
CC      The proteins can be used to screen compounds in rational drug discovery
CC      programmes. The antisense nucleic acid sequence is also useful to screen
CC      for homologous nucleic acids which are required for cell proliferation in
CC      a wide variety of organisms. The present sequence represents an
CC      essential prokaryotic cellular proliferation protein.
CC      Note: The sequence data for this patent did not form part
CC      of the printed specification, but was obtained in electronic
CC      format directly from WIPO at
XX      ftp.wipo.int/pub/published_pct_sequences.
XX
SQ      Sequence    932 AA:
Query Match          16.2%; Score 500.5; DB 22; Length 932;
Best Local Similarity 27.6%; Pred. No. 4e-19;
Matches 160; Conservative 96; Mismatches 212; Indels 111; Gaps 24;
QY      42 GIEKSEDETESTTVNDENAEATFLQKTPODNNHLEBEVKESSVESNMSSTDIAQOPSH 101
        | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db       45 gneaaaehtngelshqskett-----apen-----ktlekxdsrqq--- 82
QY      102 TTINEESQOTSDNEDSHVDSPANSIKSNTESGKEENTTIEOPNKVKVEDSTOSPQSY 161
        | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db       83 -----nlnegstsd--gpkwnesntevke--tteep----gnltsqpkdq 122
QY      162 TNIDEKISNODELMLPINIYENKARPLISTSAOPSIK-----RYTVNOOLA-EGGSNVN 215
        | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db       123 nn--damandk--nlaaqnistgaakvdstpckttllprtlmnaavltvaapggatnn 177
QY      216 HLIVKTDOSI-----TGYSODESEVIKAHADEANIYDVTFEVDKDKVSGDPM 262
        | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db       178 dkvnhtcndaidkxhvnmktgtetfawtsdvtk-----lkanyclidssveagdcf 229
QY      263 TVWDIDKNVTPSDLDTSFTRPKIKDSNGELIIATGYDMKNKKOITYFTFYDVYDKENIKAH 322
        | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db       230 tfkyygyffpgsvrlpsrqbnlynaggnlakiglydsetcttlyfftnvdqytlnsgsf 289
QY      323 KLTSTIDSKSVPNNMTKLDEVKRTALSSVNKTITVEYORPNERNRANQSMETNIDTKNH 382
        | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db       230 egvatakenatldktlaypmewtlgndkyksknlvdy--gnqkgqqlsstnylnnedm 346
QY      383 TVEQRIYIN-PLRYAKERTNV-NISGNGEGESTIIDDSFTIRIKVVYGVGDQN-----LPD 434

```

[illegible]

CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX
S0 Sequence 932 AA;

Query Match 16.2%; Score 500.5; DB 22; Length 932;
Best Local Similarity 27.6%; Pred. No. 4e-19;
Matches 160; Conservative 96; Mismatches 212; Indels 111; Gaps 24;

QY 42 GIEKSEPRTESTTVDENEAFLQKTDQDNHLEEEVKSESSVSSNSISDTQAQSH 101
Db gheakaahetngelngsknetl---apsen-----kitekvdseq--- 82
QY 102 TTINEESVQNSDVNEDSHVDFANSKIKESNTESGKEENTIEOPNKVKEDSTQSPGQY 161
Db 83 -----nmegstsd--qpkvnesontbke--ttee-----qntscqpkq 122
QY 162 TNIDEKISNODELNLPINEYENKARPLSTSAQSIK-----RYTVNQLAA--EAGSNVN 215
Db 123 nn--damankd--llaagqlstgqkdxstcpkctlliprlnmaavntvaapqgltavn 177
QY 216 HLIKVTDGSI-----TEGYDSEGVIAHAHAENLIYDTEVDDKVGSGDTM 262
Db 178 dkvhftldialdkghvkkctgntefwatsadqvl-----lkaanytlidssvkegdlf 229
QY 263 TVYDIDKNVPSDLNDSFIIPKIKDSGSELIATGTVDNKNKQITTFYDVKYENIKAKHL 322
Db 230 flkygqyfirpsavrlpsqtqnlynaqgnllakgilydsesctylctlnyvdytlnsgsf 289
QY 323 KLTYSIDSKVYNNNTKLDVEYKTLALSSVYKTIYVEYORPMENRRYANLQSMFTNIDTKNH 382
Db 290 egvatakrenatldktaypmewtlndkyskhyvdy---gnqksgqlsstenylmnedl 346
QY 383 TVEQTIYIN-PIRYAKETTNV-NISGNDEGSTIIDSTLIKVKYVGNQN-----LPD 434
Db 347 srnmvnyvnpkkyctkctefvnlc-----gykfnpdaeknfkiyev-tnqngfvdsfcpd 400
QY 435 SNRIYDSEYEDVNTDDAOLGNNDVNIINFG--NIDSYIIKVIKSKYDPNKDDVTITIQ 492
Db 401 tskldvcdkfklt---ysndkktatvdllngqsssdqyiliqvaaypnasctdngkldy 457
QY 493 TWTMOGTINNEYGEPRFASVDNTIAFSTSSGOGGDDLPEPKTYKIGDYVMEVDYDKDIGIN 552
Db 458 tletngngsswn-----sysnvngsstangd-----qkynlgdyvmedtnkdqgd 505
QY 553 TNDNEKPLSNVLVTLPDGTIS-KSVRTDEDKQYFDGV 550
Db 506 a--nekglkyvyvllkdsngkeldrtttidengkygftgl 542

RESULT 14
AAU34283
AAU34283 standard; Protein; 841 AA.
AC AU34283;
DT 14-FEB-2002 (first entry)
DE Staphylococcus aureus cellular proliferation protein #559.
XX
XX
KM Antisense: prokaryotic cellular proliferation protein;
XX antibiotic; antibacterial; drug design.
OS Staphylococcus aureus.
XX
XX
PN WO200170955-A2.
PD 27-SEP-2001.
XX
XX
PF 21-MAR-2001; 2001WO-US09180.

CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 841 AA;

Query Match 14.8%; Score 456; DB 22; Length 841;
Best Local Similarity 27.1%; Pred. No. 9e-17;
Matches 168; Conservative 94; Mismatches 222; Indels 108; Gaps 25;

```
QY 42 GIEKRSDETESTNVDENEAFLQKTPQDNTHLTEEVKSSSVES-----SNSSIDTAAQ 97
DB 45 gheakaehtngelngskneet-----apsen-----kthkvdsrcqlxhntqtatad 92
QY 98 QPSHTTIRRESVQTSQDVNEDSHVSDPANSKIKESNTESGKEINTIEQPNKVKEDSTSO 157
DB 93 qpkvt-----msdatvketssmgs-----pgnatangstlk 125
QY 158 PSGYTNIDEX---ISNOELNLPINEYENKARPLSTSAOPSIK----RYTVNQLAA- 208
DB 126 tsnvtndkssttysnecdksnl-----tqakdvstcpkttlikprlnrmavntvaap 179
QY 209 EOGSNVNHILIKVT--DOSITEGY--DDESGVTK--AHDAENLIYDVTFEYVDKVKSGDIWT 263
DB 180 qggtvndkvhnsfldaldkgvhnqgtcgtetfatsadvlikanytldsvkegdtft 239
QY 264 VDIDKNYPSDLTDSFTIPKIKDNGSEIIATGTVDNKNKQIYTFPTDYVDKXENIKAHLK 323
DB 240 fkyggyfipgsavrlpsqcnalynagqnlakglydstctntcttlytfnyvdqylnvrgsf 299
QY 324 LTSYIDKSKVPNNNTKLDVEYKRTALSSVKNRTTVEYQRPNNENFANLQSMFTNIDTKNHT 383
DB 300 qvafakrkhattdktaykmevclgndtyseeilvdy---gnkkaqpllastnynimedls 356
QY 384 VEQTIYIN-PLKYSAKERIV-NISGNGDEGSIITIDSTIIKYYKVGDNQNL-----PDSN 436
DB 357 rmtlayvnpknlcytkqtlftvnlc-----gykfnpnaknfklyevldgnqfvdstfpdls 411
QY 437 RIYDSEYEDVNDQYAOAGNNNDVINFGNIDS--PYIIKVI-----SKYDPNKDDYTT 489
DB 412 Kikdvldqfdvl---ysndkcatvaimkqcsnakyllqyaypdmnsdngkldy--- 466
QY 490 IQGTVMQTTINEYTGEEFTASIDNTIAFSTSGQGQGDLPPEKTYKIGDYWEDVDKDG 549
DB 467 -----tltdcttkysw---snysvngsstandg-----qkynlgywedtnkdg 511
QY 550 IONTNDNEKPLSNLVLTLYPDGTS-KSVRTDEDEGKYQFDGV 590
DB 512 kgda--nekgyikgyvylkdsngkeldrtctdengkyqfgy 551
```

Search completed: July 30, 2002, 10:30:30
Job time: 100 sec

Tue Jul 30 11:46:38 2002

us-09-147-405-11.ra1

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 30, 2002, 10:28:55 ; Search time 33.65 Seconds

(without alignments)
430.442 Million cell updates/sec

Title: 1 US-09-147-405-11

Perfect score: 3087
Sequence: 1 HHHHHPPSSDEEKNDYINN.....SKSVRDEDEGKYQPDGVQVD 593

Scoring table:
BLOSUM62
Gapop 10.0 , Capext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents, AA:*

- 1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
- 2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
- 3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
- 4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
- 5: /cgn2_6/ptodata/2/1aa/PCTUS.COMB.pep:*
- 6: /cgn2_6/ptodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	452	14.6	933	3	US-08-293-728-2
2	452	14.6	933	4	US-09-421-868-2
3	349	11.3	345	4	US-08-856-253-7
4	259.5	8.4	1112	2	US-08-714-402-2
5	259.5	8.4	1161	4	US-09-327-536-2
6	209	6.8	630	4	US-08-973-462-9
7	198	6.4	783	6	5231168-2
8	197	6.4	2314	4	US-09-268-347-49
9	196.5	6.1	1786	4	US-08-973-462-8
10	189.5	6.1	1664	2	US-08-642-846-2
11	189.5	6.1	1664	2	US-09-264-604-2
12	188.5	6.1	1435	2	US-08-568-459A-4
13	188.5	6.1	1435	2	US-08-487-826B-4
14	186.5	6.0	3111	2	US-08-460-309-4
15	186.5	6.0	3111	2	US-08-125-077-4
16	185	6.0	2048	4	US-09-268-347-48
17	184.5	6.0	2391	2	US-08-446-855A-2
18	184.5	6.0	2391	4	US-09-150-741-2
19	177.5	5.7	1087	1	US-08-264-002-5
20	176	5.7	3788	4	US-09-336-447A-76
21	175.5	5.7	1104	4	US-09-268-347-28
22	175.5	5.7	1104	4	US-09-268-347-34
23	175.5	5.7	1183	2	US-08-447-031A-2
24	175	5.7	1147	3	US-08-470-260-5
25	175	5.7	1147	3	US-08-471-491-5
26	175	5.7	1147	3	US-08-466-662-5
27	175	5.7	3289	2	US-08-477-451-2

28	173.5	5.6	1231	3	US-08-904-263A-4	Sequence
29	173.5	5.6	2123	4	US-08-968-685A-10	Sequence
30	166.5	5.4	984	1	US-08-257-073-3	Sequence
31	166.5	5.4	984	2	US-08-184-009-120	Sequence
32	166.5	5.4	984	2	US-08-458-356-120	Sequence
33	166.5	5.4	984	4	US-08-460-726-120	Sequence
34	166.5	5.4	1164	4	US-08-923-992A-10	Sequence
35	165	5.3	3135	1	US-08-323-110B-2	Sequence
36	165	5.3	3135	4	US-08-954-441-2	Sequence
37	164.5	5.3	1164	4	US-08-923-992A-2	Sequence
38	164.5	5.3	1420	2	US-08-540-804-14	Sequence
39	164.5	5.3	1420	2	US-08-218-265-14	Sequence
40	164.5	5.3	1420	3	US-08-521-872-14	Sequence
41	164.5	5.3	1420	4	US-08-590-399-14	Sequence
42	163.5	5.3	1098	4	US-08-923-992A-8	Sequence
43	163.5	5.3	1128	4	US-08-923-992A-6	Sequence
44	162.5	5.3	740	1	US-08-257-073-5	Sequence
45	162.5	5.3	1964	2	US-08-790-912-3	Sequence

ALIGNMENTS

RESULT 1
US-08-293-728-2
Sequence 2, Application US/08293728D
Patent No. 6008341
GENERAL INFORMATION:
APPLICANT: Foster, Timothy J.
TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene
FILE REFERENCE: 05344.105011
CURRENT APPLICATION NUMBER: US/08/293,728D
CURRENT FILING DATE: 1994-08-22
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 933
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-08-293-728-2

Query Match	14.6%	Score 452	DB 3	Length 933
Best Local Similarity	26.6%	Pred 1.6e-18		
Matches 154	Conservative 105	Mismatches 234	Indels 86	Gaps 26
QY	8	SSDEEKNDYINNOSINTDDNOI	IKKEETNNYDIEKRSDETESTVNDENATFLQK 67	
DB	51	SNESKSNDSVSAARKTDITNV	-----SDTKTSNTN---NGETSAVO 91	
QY	68	TPQDNTHLTEBEYKSSVSSNS	SDTIAQPSHTTINNEESVQTSNVEDSDVSDFANS 127	
DB	92	NP-----AQOETVOSSSNAT	TEETPVGEATTTTNOANPATQOS-SMNAEELVN- 143	
QY	128	KIKESNTESGKEBTEIOPNK	VEDSTGOPSQTINDEKISNOELNL-LPINEYENKA 186	
DB	144	--QTSNETTFNDITNV-----	SSVNSPONSIAENVSTQDITSTARPEN---NES 189	
QY	187	RPLSTSAOPSIKRYTN-----	OLAEGSNVNLIKYTDQ--SITEGYDSEGV 235	
DB	190	APGSTASKNKDVYNAVNS	SAPRMAFSLAAVAADPAAGTDITNLTQVIGI-DSGTT 248	
QY	236	IKAHDAENLIYDTEFEYDK	VSGDITMYDIDKNYPSDITDSFTLPKIKNSGELIANG 295	
DB	249	VYFHQAGYKLVNGFVPS	SAVKGDTFKITVPKELNLNGVTSRAKVPPIAAGD-QYLANG 307	
QY	296	TYNKKKQIITYFTD	VDVYKENTKAHLKLSYIDKSKVPNNKTL-DVEYKTALS--VN 352	
DB	308	VIDS--DGNVITYFTT	DDVKAFLMPAYD---DENVKKTSNVLATIGSTTAN 362	
QY	353	KITVEYQRPENRNT	ANLQSMFTNIDTKNHTVEQITVINPL--RYSAKETNWNISGNGDE 410	

```

Db 363 KTVLVDEYKGFYNLISIKGIDIDQIDKTNNTYRQTIIVNPSGDNVIAPIVLGNLKRNTDS 422
QY 411 GSTIIDDSTIKYKKGDNONLPSNRITDYSEEDVNDVAQLGNNDVNIENG- 466
Db 423 NALIDQONTSIKVKYKDNADLSESYFV-NPENFEDVTNSVNTFPNPQYKEEFTPPD 481
QY 467 NIDSPYIKYISKYDPN-KDDYTTIOQTYMTQTTINEYGE--FRTASDNTIARSTSSG 523
Db 482 QITPTIIVVNGHIDPNSGID-----LALRSTLYGYSNIITWFSMSWMDNEVAFNNGSG 534
QY 524 QGOG-DLP--PEKTYIGDY--VMEDVDKD-GIONTNDN 556
Db 535 SCBGIDKPYVPPEQDPDEGEIEPIEDSDSDPSGSDSDS 573

```

```

RESULT 2
US-09-421-868-2
; Sequence 2, Application US/09421868
; Patent No. 6177084
GENERAL INFORMATION:
APPLICANT: Foster, Timothy J.
APPLICANT: McDevitt, Damien L.
TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene
FILE REFERENCE: 05344.105011
CURRENT APPLICATION NUMBER: US/09/421,868
CURRENT FILING DATE: 1999-10-19
PRIOR APPLICATION NUMBER: 08/293,728
PRIOR FILING DATE: 1994-08-22
NUMBER OF SEQ ID NOS: 20
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 2
LENGTH: 933
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-421-868-2

```

Query Match 14.6%; Score 452; DB 4; Length 933;

Best Local Similarity 26.6%; Pred. No. 1.6e-18; Matches 154; Conservative 105; Mismatches 234; Indels 86; Gaps 26;

```

QY 8 SSDEKNDVYNNQNTDNNQIIEKTEPNYDIEKRSDEKSTSTVNDENAEFLQK 67
Db 51 SNESKNDSSSVSAAPKTDITV-----SDPKTSNTN--NGETSVNQ 91
QY 68 TPQDNHLEEEKSESSSSSIDTQAQPSHTTINEESVQSDNEDSHVDFANS 127
Db 92 NP-----AQETQSSSTNATTEETPTVGEATTTTQANPATTQS--SNTNAELVN- 143
QY 128 KIKESNTESGKEENTIEQPNKYEDSTQPSGYTINIDEKISNOBELN-LPINEYENKA 186
Db 144 --QTSMETTFNDITV-----SSVSPONSTNAEAVSSTQDTSTATISN--NES 189
QY 187 RPLSTTSAPQSIKRYTVN-----QLAEOGSNVNHLIYVTD--SITEGYDSECV 235
Db 190 APOSTDASKNDVAVNOAVNTSAPRMRASLAAVAADAPAGCTITNQLTNVTVGI-DSQTT 248
QY 236 IKAHAENLIYDVFEDVDKVKSGDMYDIDKNTVPSDLTSPITPKIKDNGSELIATG 295
Db 249 VTPHQAGYKLVKNGSVSPNSAVKGDFTKTVPEKELNLCVSTAKVPPIMAGD-QVLANG 307
QY 296 TYDNKKQITTYTVDVYENIKAKHLKLSYIDSKSVNNNTKL-DVEYKTALSS--VN 352
Db 308 VIDS-DGNVIYTFDVTNKKDVKATLTPAID----PENKKGAVNLATLIGISTTAN 362
QY 353 KTIIVYQRPENRFTANLQSMFTNIDTKNHYEQTIYINPL--RYSAKETNVNIGNGDE 410
Db 363 KTVLVDEYKGFYNLISIKGIDIDKTNNTYRQTIIVNPSGDNVIAPIVLGNLKRNTDS 422
QY 411 GSTIIDDSTIKYKKGDNONLPSNRITDYSEEDVNDVAQLGNNDVNIENG- 466
Db 423 NALIDQONTSIKVKYKDNADLSESYFV-NPENFEDVTNSVNTFPNPQYKEEFTPPD 481

```

```

QY 467 NIDSPYIKYISKYDPN-KDDYTTIOQTYMTQTTINEYGE--FRTASDNTIARSTSSG 523
Db 482 QITPTIIVVNGHIDPNSGID-----LALRSTLYGYSNIITWFSMSWMDNEVAFNNGSG 534
QY 524 QGOG-DLP--PEKTYIGDY--VMEDVDKD-GIONTNDN 556
Db 535 SCBGIDKPYVPPEQDPDEGEIEPIEDSDSDPSGSDSDS 573

```

```

RESULT 3
US-08-856-253-7
; Sequence 7, Application US/08856253
; Patent No. 6288214
GENERAL INFORMATION:
APPLICANT: Hook, Magnus
APPLICANT: Patil, Joseph M.
APPLICANT: House-Pompeo, Karen
APPLICANT: Schanam, Narayana
APPLICANT: Symersky, Jindrich
TITLE OF INVENTION: COLLAGEN BINDING PROTEIN COMPOSITIONS
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESSES:
ADDRESS: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: U.S.
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/856,253
FILING DATE: Concurrently Herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/017,678
FILING DATE: 16-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: TAWK:193
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 345 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-856-253-7

```

Query Match 11.3%; Score 349; DB 4; Length 345;

Best Local Similarity 30.8%; Pred. No. 3.7e-13; Matches 102; Conservative 61; Mismatches 136; Indels 32; Gaps 15;

```

QY 220 VTDQ--SITEGVDSEGVAKHADAENLIYDVFEVDKVKSSDPTMTVDIDKNTVPSDLTD 277
Db 25 ITNQLTNVTVGI-DSGTVVTPHQAGYKLVKNGSVSPNSAVKGDFTKTVPEKELNLCVTS 83
QY 278 SEFTIPKINDSGEIIATGYDNKNQIYTFDYDYDKYENIKAKHLKLSYIDSKSVNPN 337
Db 84 TAKVPPIMAGD-QVLANGIDS-DGNVIYTFDVTNKKDVKATLTPAID----PENV 137
QY 338 TKL-DVEYKTALSS--VKNITVEYQRPENRFTANLQSMFTNIDTKNHYEQTIYINPL- 393
Db 138 KKTGAVNLATLIGISTANKTVLVDEYKGFYNLISIKGIDIDKTNNTYRQTIIVNPSG 197

```

```

QY 394 -RYSAKETNNVNSGNGDEGSIITIDSTIIKRYKVDGNLNDPDKNRYDYSEEDYNDY 452
DB 198 DNVIAFLVLTGMLKRPDTSNALIDQNTSIIKRYKDNADLSESYFV-PPENEDYNTSVN 256
QY 453 AQLGNNNDVNINFG---NIDSPYIIKRYISKYDPN-KDYTTTIOQVMTQTTINEYGE- 506
DB 257 ITFPNPNQKVEFNTPPDQITTPYIVVYNGHIDPRSKD-----LALRSLIYQNSNI 309
QY 507 -FRTASYDNITAFSTSSGQOG-DLP--PEK 533
DB 310 IWRSMXMDNEVAFPNNGSGSGDIDKPYVPEQ 340

RESULT 4
US-08-714-402-2
Sequence 2, Application US/08714402
Patent No. 5910441
GENERAL INFORMATION:
APPLICANT: ROCHA, Claudia
APPLICANT: FISCHETTI, Vincent A.
TITLE OF INVENTION: FIBRONECTIN AND FIBRINOGEN BINDING
TITLE OF INVENTION: PROTEIN FROM GROUP A STREPTOCOCCI
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SNECKER & MATTHIS
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/714,402
FILING DATE: 16-SEP-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 016921-097
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1112 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-714-402-2

Query Match 8.4%; Score 259.5; DB 2; Length 1112;
Best Local Similarity 21.2%; Pred. No. 2.3e-07;
Matches 116; Conservative 102; Mismatches 205; Indels 123; Gaps 23;

QY 132 SNTESGKEENTIEOPKVKEDSTTSOPS-----GYTINDEKISNODELL--NLPIN 180
DB 43 ANKQSGFEIKKVDONKPKLPAGATSSLSKDGKTSVOSTSNDKGIYDAQNOLPGTYTLK 102
QY 181 E-----YENKARPLSTSAQPSIKRYTVN---QLAAEGS-NVNHLLIKVTD--OSTIE 227
DB 103 EEARPDGYDKTSRTMTVTVYENGTYKLVENPYNGEIIKSKAGSDVSSSLQLEPKMSVYS 162
QY 228 GYDDSE---GVIAHAENLIYVTEVEVDKYS-----GDPTVVIDIKNTVPSDLTDSF 279
DB 163 KYGKTEVSSGAADFYRNHAAVYFKMSFELKOKDKSETINPGDTFVLQIDRLRLNPKGISQD- 221
QY 280 TIRK-IKDNNGEIIATGTYNKKKQIYTFDVIDKYENIKAHLLKLSYIDKSKVPPNNT 338

```

```

DB 222 -IPKIIIDANSPLAIGKHYHNEHOLITFTFDIAGLDKVLASLSELENKTEYLEN- 278
QY 339 KLDVEYKTALSSVYKTI-----TVEYORPNE-----NRTANLOSMTNIDTK 380
DB 279 -----TSISNFKSTIGGEITYKGVNVLVGNESMTESNYITNGLSNVGSGIESYVTE 331
QY 381 NHTVEQTIYINPLKRYSAKETNNVNSGNG-----DEGSIITIDSTIIKRYKVDGN 430
DB 332 TGEFVWYVYVNPNTNIPYATNMILGFEGRANSNTSDELDNANTSAGELIQVEEPEGE 391
QY 431 NLPSDNRIYDYSEXEDVT-----NDYVQLGNNNDV---NINFGN--IDSPYIIKRYSK 479
DB 392 KLPSSYGV-----DYKTLRLRDITAGLNGPQMTKRQNDGNONNQNAPFIKRYGK 444
QY 480 YDPNKDYYTIIQVYMTQTTINEYTGERTAS-----YDNTIAFSTSSGQOG--- 527
DB 445 TDOS-----GKPLVQSNLASFEGASEYAAFTPVGQVYFONEIALSPKSGSGSKSE 497
QY 528 ----DLPPEKRYKIGDYWEDVDKGIQNTNDNEKLSNVLVTLTPDGTSKSVR--TDE 581
DB 498 FTKPSITVANIUKRAQLRKKMSTDNV-----PLPEAAFEKSSNGNSQKLEASSVT 549
QY 582 DKYQF 587
DB 550 QGEVHF 555

RESULT 5
US-09-327-536-2
Sequence 2, Application US/09327536
Patent No. 6355477
GENERAL INFORMATION:
APPLICANT: FISCHETTI, Vincent A.
APPLICANT: ROCHA, Claudia
TITLE OF INVENTION: FIBRONECTIN AND FIBRINOGEN BINDING PROTEIN FROM GROUP A
TITLE OF INVENTION: STREPTOCOCCI
FILE REFERENCE: 022927-008
CURRENT APPLICATION NUMBER: US/09/327,536
CURRENT FILING DATE: 1999-06-08
PRIOR APPLICATION NUMBER: US 08/714,402
PRIOR FILING DATE: 1996-09-16
NUMBER OF SEQ ID NOS: 2
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 2
LENGTH: 1161
TYPE: PRT
ORGANISM: SFFBP gene
US-09-327-536-2

Query Match 8.4%; Score 259.5; DB 4; Length 1161;
Best Local Similarity 21.2%; Pred. No. 2.4e-07;
Matches 116; Conservative 102; Mismatches 205; Indels 123; Gaps 23;

QY 132 SNTESGKEENTIEOPKVKEDSTTSOPS-----GYTINDEKISNODELL--NLPIN 180
DB 43 ANKQSGFEIKKVDONKPKLPAGATSSLSKDGKTSVOSTSNDKGIYDAQNOLPGTYTLK 102
QY 181 E-----YENKARPLSTSAQPSIKRYTVN---QLAAEGS-NVNHLLIKVTD--OSTIE 227
DB 103 EEARPDGYDKTSRTMTVTVYENGTYKLVENPYNGEIIKSKAGSDVSSSLQLEPKMSVYS 162
QY 228 GYDDSE---GVIAHAENLIYVTEVEVDKYS-----GDPTVVIDIKNTVPSDLTDSF 279
DB 163 KYGKTEVSSGAADFYRNHAAVYFKMSFELKOKDKSETINPGDTFVLQIDRLRLNPKGISQD- 221
QY 280 TIRK-IKDNNGEIIATGTYNKKKQIYTFDVIDKYENIKAHLLKLSYIDKSKVPPNNT 338
DB 222 -IPKIIIDANSPLAIGKHYHNEHOLITFTFDIAGLDKVLASLSELENKTEYLEN- 278
QY 339 KLDVEYKTALSSVYKTI-----TVEYORPNE-----NRTANLOSMTNIDTK 380

```

Db 279 -----TSISNFKSTIGGQETTYKGVNVLNGNESTKESNTITNGLSNVGSIESYNT 331
QY 381 NHVEQTIYINPLRYSAKETNNVINGNG-----DEGSTIIDSTIIKYKVDNQ 430
Db 332 TGEFVYVYVNPRTNIPATNMLMGFRARSNTSDLENDANTSALGEIQYVEPEGE 391
QY 431 NLPDSNRIDYSEYEDVT-----NDQYAOGLNNNDY-----NINEGN--IDSPYIIVISK 479
Db 392 KLPSSISGV-----DYTKILTLTDITAGLGNFQNTKORIDFNNIQKAFIIVTKG 444
QY 480 YDPNKADYTIQOTYMTQTTINEYTGEPRTAS-----YDNTIAFSTSGOGOG--- 527
Db 445 TDS-----GKPLVYOSNLASFSGASEYAAFTPVGVNVPQNEIALSPKSGSGKSE 497
QY 528 ----DLPEKTYKIGDYVEDVDKQIQNTNDNEKPLSNVLTITPPDGTSKSVR--TDE 581
Db 498 FTKPSTTVANLKRVAOLRKKMSTDNV-----PLPEAFELRSSNGNSOKLEASSNT 549
QY 582 DKRYOP 587
550 QGEVHF 555

RESULT 6
US-08-973-462-9
Sequence 9, Application US/08973462B
Patent No. 6191270
GENERAL INFORMATION:
APPLICANT: DRUILHE, PIERRE
APPLICANT: DAUBERSIES, PIERRE
TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
FILE REFERENCE: 0660-0125-0 PCT
CURRENT APPLICATION NUMBER: US/08-973,462B
CURRENT FILING DATE: 1998-02-06
EARLIER APPLICATION NUMBER: PCT/FR96/00894
EARLIER FILING DATE: 1996-06-12
EARLIER APPLICATION NUMBER: FR 95/07007
EARLIER FILING DATE: 1995-06-13
NUMBER OF SEQ ID NOS: 29
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 9
LENGTH: 630
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Polypeptide
US-08-973-462-9

Query Match 6.8%; Score 209; DB 4; Length 630;
Best Local Similarity 19.5%; Pred. No. 8.6e-05;
Matches 116; Conservative 121; Mismatches 206; Indels 152; Gaps 28;

QY 9 SDEKNDVYNN-----NQSINTDNNQIIEKETNANYDIGKESDRTST 54
Db 59 SDELNELNLNDVNGVEKENILEESQVNDIFNLSLVQOEOOHNEKEVEEEND 118
QY 55 -----TVDENE-----ATFLCK-----TQDNTHLLEEVK 81
Db 119 EESVEENVEENVEENDSDSSVASSVESIDSSIEENAVPTVEEIVAPVEIV 178
QY 82 ESSSVSSSSSIDTAQPSHTTINRESVQTSNDVEDS---HVSDFANSKIKES---NTE 135
Db 179 APSVSVSVAPSVESVARS-----VEESV--AENVEESVAVNEVEIYAPSVESVAVNE 231
QY 136 SKGKENTIEQNKYKEDSTTQSPGTYNIDKIS--MODELINLPIN-----YENKARPL 189
Db 232 ESVAENVEESVAVNEVEESVAV-----NVEESVAVNEVEIYAFVVEESVAVPTVEIYAPT 285
QY 190 STTSAPSKIRVTYNQLAEOGSNNHLIKVTDSITGCVYDSDSGVIAHAEMLIYDT 249
Db 286 VEEVAPVVEIYAPSVESVAVSV-----EESVAVNEVEESV--AENVEESVAV 331

QY 250 FEVDKRYKSG--DTMTVIDDKNTVPSDLTDSFTIPKIKDNGSEIATGTYNKKNQI--- 304
Db 332 ENVEESVAVNEVEESVAVNEVEEIVAPS--VEEIVAPVVEESVAVNEVATNLSLNLG 389
QY 305 --TYFTDVV--DKYENIKAHKLKLTSDISKYPNNNTKLDVEKTLASSV----- 351
Db 390 IETEEIKDILNEIEVEKENV--VTTLKEVE-----ETIASVTTFSILNEIOENTYT 442
QY 352 NRTIYVEXORPENRTANLQSMFTNIDTKNHVEQTIYINPLRYSAKETNNVINGNGDGC 411
Db 443 NDTIEKLELHEN--VLSALENQSEKKEVIDYEEVKEEYATLLIEVEAEDE 499
QY 412 STIIDSTIIKYV-----KVCQN--ONLPDS--NRIYDSEY-----EDVND 450
Db 500 S-----ESTITTEIFENLEENAVESNEKVAENLEKNETYENTLADKYEETVEISGESLENN 555
QY 451 D-----YAOGLNN-----NDVNNINGNDSPYIIVISKYOPNDDYTIQOTV 494
Db 556 EMDKAFSEIFDYNVAGIOENLITGMFRSITISVIOSEEKVDLNEVSSIIDNI 610

RESULT 7
5231168-2
Patent No. 5231168
APPLICANT: DIEZIEL, MORTEN; BORRE, MARTIN; JEPSEN, SOREN;
VUUST, JENS; RIENECK, KLAUS; WIND, ANNETTE; JAKOBSEN, PALLE H.
TITLE OF INVENTION: MALARIA ANTIGEN
NUMBER OF SEQUENCES: 19
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/409,658
FILING DATE: 18-SEP-1989
SEQ ID NO: 2
LENGTH: 783
5231168-2

Query Match 6.4%; Score 198; DB 6; Length 783;
Best Local Similarity 18.5%; Pred. No. 0.00049;
Matches 118; Conservative 127; Mismatches 216; Indels 176; Gaps 28;

QY 4 HHHSSDEKNDVYNNQINNDT-----DNNQIKK-----EETNNYDG 42
Db 54 HHENSSSEVSNLENENEFVESDKVTEPAHEEYVESSESNPEANENESSIEEHOETI 113
QY 43 IEKNSDRTSTVNDENEAFTLOKT-----PQDNTHLTEEVKSSSVSSSSSIDTAQ 97
Db 114 VPEQNDSESGGLVDNEGDFEEPNHFEFPDQ---DSBLSENEIYVESKVSSEPAE 169
QY 98 QPSHTTINRESV--QTSNDVEDSHVDFANSKIKESNTESKRENTIEQPKVKEDSTS 156
Db 170 ---HVEIYSEKVSSEPAHE--IVSEKSTSEPAE--HVESVSEOSN--NEPSEKKGVPVS 222
QY 157 QPSGTYNIDKISMODELINLPINEY--NKARPLSTTSAPSKIRVTYNQLAEOGSNNV 215
Db 223 KP--FEEL--EKVDQPKIVDQIIEPNVDQPNQVEPVSF--VKEKPESEKNKIAS 277
QY 216 HLKVTDSITGIDDSGCVIAHDAENLIYDVFVEDVKYKSGDTMTVIDDKNTVPSDL 275
Db 278 -----VPEVEKEKENVSEVEEKNQSQSVSEIYVN--EDEREVDYHTQLDDHKTVDPEI 331
QY 276 TDSFTIPKIKDNGSEIATGTYNKKNQIYTFDIDYDKIENIAHLKLTISYIDSKYVN 335
Db 332 VEVEIIPS-----ELHNEVAVHEIYVEIEVEFEPN 362
QY 336 NNTKLDEYKTLALSVKNTIYVQRPNE--NRTANLQSMFTNI-----DTRKNTVEQTI 388
Db 363 QNN-----EFOEINEDDKSAHIEIYVEIEIYEDDKNKEVHEI 403
QY 389 YINPLRYSAKETNNVINGNGDGSTIIDSTIIYKVGNDQNPDSNR---IYDSEY 444
Db 404 -----VEVEIILPEDKNEKGOHEIYVEIEI---LPEDDKNKEVHEIYEV 445
QY 445 EDVTNDQYAOGLNNNDVNNINGNDSPYIIVISKYOPNDDYTIQOTYMTQTTINEYT 504

QY 491 -----OQVYMTOTINEXYGERFASDNTAFSTSG----- 523
DB 1053 DNENKKEGLINKLENISSTEGVETVEHEQNVYVDVFPAMKQFGLINAEGLKE 1112
QY 524 -----OGGDDLPPEKTYKIGDYVWEDVDKQIONTNDNEKPLSNVL-----VTLT 568
DB 1113 MFENLEDFVSESDVITVEIK-DEPYQKREKEVETIIEKEMENYDVLEEKEDLTOK 1171
QY 569 YPDGTSKSVRTDEGKYQPDGVQ 591
DB 1172 MDAVESIEISSDSKEETESIK 1194
RESULT 10
US-08-642-846-2
Sequence 2, Application US/08642846
Patent No. 5886151
GENERAL INFORMATION:
APPLICANT: HOSTETTER, MARGARET K.
APPLICANT: GALE, CHERYL A.
APPLICANT: BENDEL, CATHERINE M.
APPLICANT: TAO, NIAN-JUN
TITLE OF INVENTION: CANDIDA ALBICANS GENE, INTEGRIN-LIKE
TITLE OF INVENTION: PROTEIN, ANTIBODIES, AND METHODS OF USE
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESSES:
ADDRESSEE: MUEITING, RAASCH, GEBHARDT & SCHWAPPACH, P.A.
STREET: 119 NORTH FOURTH STREET, SUITE 203
CITY: MINNEAPOLIS
STATE: MINNESOTA
COUNTRY: USA
ZIP: 55401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/642,846
FILING DATE: 03-MAY-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: MUEITING, ANN M.
REGISTRATION NUMBER: 33,977
REFERENCE/DOCKET NUMBER: 110.00280101
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-305-1217
TELEFAX: 612-305-1228
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1664 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-642-846-2
Query Match 6.1%; Score 189.5; DB 2; Length 1664;
Best Local Similarity 20.4%; Pred. No. 0.0039;
Matches 131; Conservative 105; Mismatches 236; Indels 171; Gaps 31;

DB 248 EV-----NSREPALTDKMLKRENFSLDEKVALYLSPTNNNNKNSVSDMHLQ-----N 299
QY 156 SPSGYTINDEKISNODELLNLPINEYENKARL-STSAQPSIKRYTVNQLAEQGSNV 214
DB 300 LODASKNKTNENHNSFALKAPKNDIEN---PLNSLTYNADISLRSSGSSQSLQSLRND 356
QY 215 NMLIKV---TDOSITGYDSEGVIRAKDAENLIYDTEFVDDKYSQGTMTYDIDKNY 271
DB 357 NRYLEVPSPSKVNPGLSLNDIGKFS-----EVEESLT 392
QY 272 PSDLT-DSEFTIPKID-----NSGEIANGTYDNKKQIYTFYDVKYENIKAH----- 321
DB 393 PRDLSRDKLETTKEHDAPENHNENFIDAKSTYNNKG-QLYVSSDDHLDSDRBYNTEGS 451
QY 322 -LKTISYIDSKVYNN-----NTLIDVEYKTALSSVNK 353
DB 452 ILNLTNSASQSLNALEROROTOQAAPPEETSFSFNINIVKQEPKSNLEFVYV 511
QY 354 TIVEYQRENENTANLQSMFTNIDTKNH-TVEQTIYINPLRYSAKETVYNVISG----- 406
DB 512 TIKPEVSAITEIKAPKREFSSRLRIKNEDEIPEPADIHPPKKNENANSHVEDTDLAKRA 571
QY 407 -NGDESGTIIIDSTIIKY-YKVGDNQNPDSNRIDYSEVEDYTNDDYQAGNN---NDV 461
DB 572 LNDESDDTQNSTKMSIRPHIDSDMKLEDSND---GRED--NDISRFKSDILNDV 625
QY 462 N-----INFGNIDSPYIIKYSKIDPNKDYTTIQQVYMTOTINEXYGER-----T 509
DB 626 SQTSDIIGKYGNSSEIITKTJLA---PPRSDNDKXNSKSLDPPANNESLQOOLEVPH 682
QY 510 ASYDNTIARFSTSGQOGDLPPEK---TYKIDGY-VWEDVDK 547
DB 683 KEDDSILANSSNIA-----PPEELTPVYEANDYSSFNDYTK 719
RESULT 11
US-09-264-604-2
Sequence 2, Application US/09264604
Patent No. 6346411
GENERAL INFORMATION:
APPLICANT: HOSTETTER, MARGARET K.
APPLICANT: GALE, CHERYL A.
APPLICANT: BENDEL, CATHERINE M.
APPLICANT: TAO, NIAN-JUN
TITLE OF INVENTION: CANDIDA ALBICANS GENE, INTEGRIN-LIKE
TITLE OF INVENTION: PROTEIN, ANTIBODIES, AND METHODS OF USE
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESSES:
ADDRESSEE: MUEITING, RAASCH, GEBHARDT & SCHWAPPACH, P.A.
STREET: 119 NORTH FOURTH STREET, SUITE 203
CITY: MINNEAPOLIS
STATE: MINNESOTA
COUNTRY: USA
ZIP: 55401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/264,604
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/642,846
FILING DATE: 03-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: MUEITING, ANN M.
REGISTRATION NUMBER: 33,977
REFERENCE/DOCKET NUMBER: 110.00280101
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-305-1217

TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 3111 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-460-309-4

Query Match 6.0%; Score 186.5; DB 2; Length 3111;
Best Local Similarity 19.9%; Pred. No. 0.013;
Matches 142; Conservative 102; Mismatches 256; Indels 215; Gaps 32;

28 NNOIIRKEET-----NNDIGIKRSEDRTSTTNDENATFLQTPDNTLITEE 78
1694 NERAIKINLETIGRDEAFERNLEGLQK-----EIDQMIKELRKKNLETOKEIAED 1743
79 EVKSSSVSSNSIDTAQPSHTTIN---REESVQSDVSDS-HVSDRANKIKESN- 133
1744 ELVAABALLKKYKKLFGESRGENEEMEKDLREKLADYKKKVDAMDLLREATDKIREANR 1803
134 -----TESGKE--ENTIEOPNKV-----KEDSTTS 156
1804 LFAVNOKNMTALEKKKEAVESGKRQIENLTKEGNDILDEANRLADEINSIIDVEDIQTK 1863
157 QPSGYTNIDEKISN---ODELLNPINYEYENKARPLSTSA-----QPSIKRVYVNO 206
1864 LPPMSEELNDKIDDLGSEIKDRKLAESVQAEASHAQLNDSAVLDGILDEAKNISFNAT 1923
207 AAEQG-SNVNHLIKYVDQSTIEGVDSEGYIKAHDAENLIYDVFEDVKVSGDMTVD 265
1924 AAFKAVSNIKDYI-----DEAEKVAK--EAKDLAHEAT-----KLATGPR---G 1962
266 IDKNTVPSDLTDSFTI-----PRKIDNSGEIATGT---YLNKKQIYTFPTDYVD 313
1963 LKEDAKGCLQKSFRIINEAKKLANDYKEDHNLGKTRIEVALRNGDILFTLNDTLG 2022
314 KYENI--KAHLKLTSLYIDSKVNNNTK-----LDVEYKTAALSSVKNKTT 356
2023 KLSAIPNDTAALKQAVVAKRQANDTAKDYLQITELHONLDGLKKNYKRLADSVAKTNA 2082
357 VERQRPENNR-TANLOSMTFNID-----TKNHTVEQTIYINPLRYS 396
2083 V-VKDSKKNITADADATYKNELEQADRLDKPIKELEDNKKNISEIKELINQARKQ 2141
397 AKETNNISGND-----EGS-----TIIDSTIIVK-----YK 425
2142 ANSIVSVSSGDCIRTYKPEIKKGSYNNIVVWKTAVADNLLFYLSAKFIDFLATEMR 2201
426 VGDNOULPDSNRILYDSEEDVTND-----YAOIGNNNDYVINEGNDSPYIIVIS 478
2202 KGVSVSLMDVSGGVGEVYEDLTIDDSYWRIVASPTGRNGTISVR--ALDGRKASTVPS 2259
479 -KIDPKKDQVTTIQVYTMQTTINETYGERTASDYDTIAFSISSQGGQDLPPEKTKYI 537
2260 THHSTSPPGTITLDVANAMLFVGLGLTKLKKDAVRVITFTGCMKEITYPDNRK-1 2314
538 GDVMEVDVDKGIQNTNDNEKPLSNVLTLYTPD--GTSKSVRT-DEDKGYQFDG 589
2315 G--LNMFRKEG-----DCKGCTVSPVESESGITQFDG 2346

RESULT 15
US-08-125-077-4
Sequence 4, Application US/08125077
Patent No. 5872231 5840863
GENERAL INFORMATION:
APPLICANT: Engvall, Eva
APPLICANT: Leivo, Ilmo
TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin
TITLE OF INVENTION: Fragments and Uses Thereof

NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/125,077
FILING DATE: 22-SEP-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US PCT/US 94/10730
FILING DATE: 21-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/472,319
FILING DATE: 30-JAN-1990
APPLICATION NUMBER: US 07/919,951
FILING DATE: 27-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LA 9721
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 3111 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-125-077-4

Query Match 6.0%; Score 186.5; DB 2; Length 3111;
Best Local Similarity 19.9%; Pred. No. 0.013;
Matches 142; Conservative 102; Mismatches 256; Indels 215; Gaps 32;

28 NNOIIRKEET-----NNDIGIKRSEDRTSTTNDENATFLQTPDNTLITEE 78
1694 NERAIKINLETIGRDEAFERNLEGLQK-----EIDQMIKELRKKNLETOKEIAED 1743
79 EVKSSSVSSNSIDTAQPSHTTIN---REESVQSDVSDS-HVSDRANKIKESN- 133
1744 ELVAABALLKKYKKLFGESRGENEEMEKDLREKLADYKKKVDAMDLLREATDKIREANR 1803
134 -----TESGKE--ENTIEOPNKV-----KEDSTTS 156
1804 LFAVNOKNMTALEKKKEAVESGKRQIENLTKEGNDILDEANRLADEINSIIDVEDIQTK 1863
157 QPSGYTNIDEKISN---ODELLNPINYEYENKARPLSTSA-----QPSIKRVYVNO 206
1864 LPPMSEELNDKIDDLGSEIKDRKLAESVQAEASHAQLNDSAVLDGILDEAKNISFNAT 1923
207 AAEQG-SNVNHLIKYVDQSTIEGVDSEGYIKAHDAENLIYDVFEDVKVSGDMTVD 265
1924 AAFKAVSNIKDYI-----DEAEKVAK--EAKDLAHEAT-----KLATGPR---G 1962
266 IDKNTVPSDLTDSFTI-----PRKIDNSGEIATGT---YLNKKQIYTFPTDYVD 313
1963 LKEDAKGCLQKSFRIINEAKKLANDYKEDHNLGKTRIEVALRNGDILFTLNDTLG 2022
314 KYENI--KAHLKLTSLYIDSKVNNNTK-----LDVEYKTAALSSVKNKTT 356
2023 KLSAIPNDTAALKQAVVAKRQANDTAKDYLQITELHONLDGLKKNYKRLADSVAKTNA 2082

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 30, 2002, 10:28:55 ; Search time 52.53 Seconds

(Without alignments)
1084.732 Million cell updates/sec

Title: US-09-147-405-11

Perfect score: 3087
Sequence: 1 HHHHHHSSDDEKNDYINN.....SKSVFDEGKQFDGYVD 593

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3015	97.7	1092	2 T30214	fibrinogen-binding
2	1146.5	37.1	1166	2 T28680	fibrinogen-binding
3	1139.5	36.9	1141	2 E89824	hypothetical prote
4	576	18.7	1315	2 T28679	fibrinogen-binding
5	559	18.1	1385	2 D89824	hypothetical prote
6	490	13.9	953	2 C89824	hypothetical prote
7	452	14.6	933	2 S41539	fibrinogen-binding
8	405.5	13.1	877	2 D89852	fibrinogen-binding
9	376	12.2	940	2 F90070	clumping factor B
10	376	12.2	940	2 S19702	fibrinogen-binding
11	345.5	11.2	961	2 G90053	hypothetical prote
12	326	10.6	1038	2 H90053	hypothetical prote
13	304	9.8	1018	2 A32192	fibrinogen-binding
14	239	7.7	1039	2 T30856	protein F2 - strep
15	225	7.3	1463	2 T30290	AAS surface protei
16	222	7.2	1431	2 T18429	hypothetical prote
17	221.5	7.2	1431	2 AG3085	hypothetical prote
18	216.5	7.0	1431	2 T18501	hypothetical prote
19	214	6.9	1093	2 T18275	1-phosphatidylinos
20	213.5	6.9	1093	2 G71611	hypothetical prote
21	213	6.9	821	2 S67087	hypothetical prote
22	212.5	6.9	1125	2 E90598	membrane nuclease
23	211.5	6.9	3724	2 T18427	hypothetical prote
24	209	6.8	4550	2 T18440	hypothetical prote
25	208.5	6.8	1072	2 A86827	hypothetical prote
26	208	6.7	1127	2 T28317	ORF MSV156 hypot
27	207.5	6.7	769	2 T28676	rhodopsin protein
28	207	6.7	769	2 F89870	serine proteinase
29	207	6.7	1939	2 T18572	repeat organellar

30	206	6.7	1302	1 J06009	surface-located me
31	205.5	6.7	1365	2 T30822	Impi protein - Myc
32	205.5	6.7	2510	2 T28160	hypothetical prote
33	205	6.6	2481	2 D90011	FmB protein (limp
34	205	6.6	3844	2 T18402	asparagine/asparta
35	203	6.6	537	2 A23770	asparagine-rich pr
36	202.5	6.6	1997	2 F71607	DNA helicase II BR
37	202.5	6.6	2269	2 T28677	rhodopsin protein
38	202	6.5	1308	2 E71622	probable membrane
39	201.5	6.5	4688	2 F82885	hypothetical prote
40	201	6.5	665	2 B71609	hypothetical prote
41	200.5	6.5	3147	2 T18674	hypothetical prote
42	200	6.5	2150	2 F71629	sensory transducti
43	199.5	6.5	1139	1 E64234	cytochrome-c acc
44	198.5	6.4	807	2 B71605	hypothetical prote
45	198.5	6.4	1805	2 T02712	similar to late em

ALIGNMENTS

RESULT 1	
T30214	fibrinogen-binding protein - Staphylococcus epidermidis
C:Species: Staphylococcus epidermidis	
C:Date: 22-Oct-1999 #sequence revision 22-Oct-1999 #text change 11-May-2000	
C:Accession: T30214	
R:Nilsson, M.; Frykberg, L.; Flock, J.I.; Pel, L.; Lindberg, M.; Guse, B.	
Infect. Immun. 66, 2666-2673, 1998	
A:Title: A Fibrinogen-binding protein of Staphylococcus epidermidis.	
A:Reference number: 220781; MIMD:98261511	
A:Accession: T30214	
A>Status: preliminary; translated from GB/EMBL/DBJ	
A:Molecule type: DNA	
A:Residues: 1-1092 <NID>	
A:Cross-references: EMBL:Y17116; NID:e1296734; PID:e1296735; PIDN:CAA76638.1	
Query Match 97.7%; Score 3015; DB 2; Length 1092;	
Best Local Similarity 99.7%; Pred. No. 9,6e-122;	
Matches 582; Conservative 2; Mismatches 0; Indels 0; Gaps 0;	
QY 8 SSDEKNDYINNQSINTDNNQIKKEFNNDGIEKSEDETESTNDENAEFLQK 67	
DB 75 SSDEKNDYINNQSINTDNNQIKKEFNNDGIEKSEDETESTNDENAEFLQK 134	
QY 68 TPQDNTHLTEEEYKSSSVSSNSIDTAQOPSHTTINRESVQTSNDVSDSHVDFANS 127	
DB 135 TPQDNTHLTEEEYKSSSVSSNSIDTAQOPSHTTINRESVQTSNDVSDSHVDFANS 194	
QY 128 KTESNTESGKENTIEOPKPKVEDSTSPSGTYNIDENISNDELINPIVEYENKAR 187	
DB 195 KTESNTESGKENTIEOPKPKVEDSTSPSGTYNIDENISNDELINPIVEYENKAR 254	
QY 188 PLSTSAOPSIKRVNTQALAEQSNVHLIKYTDOSITFGYDSEGVAKAHDAENLIYD 247	
DB 255 PLSTSAOPSIKRVNTQALAEQSNVHLIKYTDOSITFGYDSEGVAKAHDAENLIYD 314	
QY 248 VFEVDVKVSGDPTMVDIDKNTVPSDLTDSFTPIKIKNSGEIATGTYDNKKOITYT 307	
DB 315 VFEVDVKVSGDPTMVDIDKNTVPSDLTDSFTPIKIKNSGEIATGTYDNKKOITYT 374	
QY 308 FTDYVDKENIKAKHLKSTYIDSKYPPNNNTKLDVEYKTAASSVNTTYVEYRPENKRT 367	
DB 375 FTDYVDKENIKAKHLKSTYIDSKYPPNNNTKLDVEYKTAASSVNTTYVEYRPENKRT 434	
QY 368 ANQSMFTNIDTKNHVEQRTYINPLRYSAKETNNVNSGNGDSPTIIDSTIIKYKYKG 427	
DB 435 ANQSMFTNIDTKNHVEQRTYINPLRYSAKETNNVNSGNGDSPTIIDSTIIKYKYKG 494	
QY 428 DNQNLDSNRIYDSEYEDVTNDYQALGNNNDVNIIFGNDSPYIIKYISKYDPKKDXY 487	
DB 495 DNQNLDSNRIYDSEYEDVTNDYQALGNNNDVNIIFGNDSPYIIKYISKYDPKKDXY 554	

Db 643 KSVRTDANGHYEGGLK 655

RESULT 4

T28679

fibrinogen-binding protein homolog - Staphylococcus aureus

C:Species: Staphylococcus aureus

C:Date: 15-Oct-1999 #sequence, revision 15-Oct-1999 #text, change 11-May-2000

C:Accession: T28679

R:Jostesson, E.; McGree, K.; N. Eldholm, D.; O'Connell, D.; Cox, J.; Hook, M.; Foster, T.

Microbiology 144, 3387-3395, 1998

A:Title: Three new members of the serine-aspartate repeat protein multigene family of Staphylococcus aureus

A:Reference number: Z20510; MUID:99096700

A:Accession: T28679

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1315 <JOS>

A:Cross-references: EMBL:AJ005646; NID:e1318791; PID:e1318792; PIDN:CAA0651.1

C:Genetics:

A:Gene: sdrc

Query Match 18.7%; Score 576; DB 2; Length 1315;

Best Local Similarity 30.1%; Pred. No. 2.4e-17;

Matches 186; Conservative 112; Mismatches 228; Indels 92; Gaps 29;

Db 8 SSDEKNDV---INNOSINTDNNQIIRKERTNYDGIKRSKEDRTSTNVDE-NEAT 63

Db 55 STNKLNATISASDNOSSDKVDMQLOEDNTKNDOKEMVSSQGNFTSNGKLIKE 114

Db 64 FLOKTPQDNTHLLEEVKESSESSNSSIDTAQPSHTTINREESVQTSNVEDSHVD 123

Db 115 SVQSTGKVKVEVSTAKSDEQASPKSTNEDLNTKQ-----TISNQGLO-PDLQEKSVVN 168

Db 124 FANSKIKESNTESGKEENTIEQPNKVEDSTSPSGTYNI-DEKISNDEL-NLPIN 180

Db 169 -----VQPTNEKKVD-----AKTEST-----LVNSDAIKSDELIVNNSMS 209

Db 181 EYENKARPLSTSAQP---SIKRYVNOALAEQSNVNLIKYTDOSITEGYDSEGYI 236

Db 210 NNNENNADILPKSTAPKRLNTRMRIAAYQPSSTEAKNVNLTITNTLVVDADKNKIV 269

Db 237 KAHDEMLIYDVFEDVKVSGDTMTVDIDKNTV-----PSDLTDSFTIKIKD-NSG 289

Db 270 PAQDYLSLKSQIT--VDDKVSQDYFTIKY--SDTVQVYGLNPEIDIKN--IGDIKDPNNG 323

Db 290 ELIATGVNKKKQITFTDYVDKYEIKAKHLKLSYIDKSKVPNNTKLDEYKTAIS 349

Db 324 ETIAAKHDTANMLITYFTDYVDKNSVQMGINSIYMDADTIP--VSKNDVEFNVITG 381

Db 350 SYNKITVEYQRPN--ENRTANLQSMFTNIDTKNHT-----VEQTIYINPLRSYSAKE 399

Db 382 NNTTKTANIQPDYVYVNEKNSIGSAFT--ETVSHVGNKNEPGYKQTIYVNSSENSLN 439

Db 400 TNVNT-----SGNDESGTIIIDSTIIKRYKVGDNQNLPSNRITYDS--EYEDVTDDA 453

Db 440 AKLKVQVHSSYPNNIGQINKVDIKIQVPRGYTL--NKGYDVNTELVETN-QYL 495

Db 454 Q---LGNNDVNVINGNINSPYIIKISKYDPKDDYTTIQQVTFMOTINTEYGEFRA 510

Db 496 QKTTYDDNNSAVIDFGNADSAIVVYVNTKQRYNSESPLVQMAITLSSIGN-----KSV 549

Db 511 SYDNTIAFTSSGQGGDLPPEKTYKIGDYVWEDVDKGIQNTNDRKPLSNLVLTLYTP 570

Db 550 STGNALGFTNNGSGAG---QEVYKIGYVWEDVTNKNVQEL--GEKGVNTVTVV-FD 602

Db 571 DGTSKSVR---TDEDGY 585

Db 603 NNTNTKVGAVTKEDGSY 620

hypothetical protein sdrc [imported] - Staphylococcus aureus (strain N315)

C:Species: Staphylococcus aureus

C:Date: 10-May-2001 #sequence, revision 10-May-2001 #text, change 22-Oct-2001

C:Accession: D89824

R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cul, L.; O

ma, A.; Mizutani, U.; Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.

C:Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.

Lancet 357, 1225-1240, 2001

A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.

A:Reference number: A89758; MUID:21311952; PMID:11418146

A:Accession: D89824

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1385 <NUR>

A:Cross-references: GB:BA000018; PID:913700454; PIDN:BAR41751.1; GSPDB:GN00149

A:Experimental source: strain N315

C:Genetics:

A:Gene: sdrc

Query Match 18.1%; Score 559; DB 2; Length 1385;

Best Local Similarity 28.8%; Pred. No. 1.4e-16;

Matches 177; Conservative 117; Mismatches 235; Indels 86; Gaps 27;

Db 8 SSDEKNDV---INNOSINTDNNQIIRKERTNYDGIKRSKEDRTSTNVDE-NEAT 63

Db 55 STNKLNATISASDNOSSDKVDMQLOEDNTKNDOKEMVSSQGNFTSNGKLIKE 114

Db 64 FLOKTPQDNTHLLEEVKESSESSNSSIDTAQPSHTTINREESVQTSNVEDSHVD 123

Db 115 SVQSTGKVKVEVSTAKSDEQASPKSTNEDLNTKQ-----TISNQGLO-PDLQEKSVVN 168

Db 124 FANSKIKESNTESGKEENTIEQPN-KVEDSTSPSGTYNI-DEKISNDEL-NLPIN 182

Db 169 -----VQPTNEKKVDKAKTESTTNVNSDAIKS-----NATLVNNSNSNN----- 211

Db 183 ENKARPLSTSAQPSIKRYVNOALAEQSNVNLIKYTDOSITEGYDSEGYI 238

Db 212 ENNADILPKSTAPKRLNTRMRIAAYQPSSTEAKNVNLTITNTLVVDADKNKIV 271

Db 239 HDAENLIYDVFEDVKVSGDTMTVDIDKNTV-----PSDLTDSFTIKIKD-NSGEI 291

Db 272 QDYLSLKSQIT--VDDKVSQDYFTIKY--SDTVQVYGLNPEIDIKN--IGDIKDPNNG 325

Db 292 IATGTYDNKKKQITFTDYVDKYEIKAKHLKLSYIDKSKVPNNTKLDEYKTAIS 351

Db 326 IATKADHTANMLITYFTDYVDKNSVQMGINSIYMDADTIPVD--KKDVPSTVITGQ 383

Db 352 NKTIVVEYQRPN--ENRTANLQSMFTNIDTKNHT-----VEQTIYINPLRSYSAKE 401

Db 384 ITTTADITTPPAKKEADNNSIGSAFT--ETVSHVGNVEDPGYKQTIYVNSSENSLN 441

Db 402 VNISNGDESGT---IIDSTIIKRYKVGDNQNLPSNRITYDS--SEYEDVTND--DYA 453

Db 442 LKVEAYHPKYPFNIQINQNTNIRIKIYRVEGYTL--NKGYDVNTELVETN-QYL 498

Db 454 QLGNNNDVNVINGNINSPYIIKISKYDPKDDYTTIQQVTFMOTINTEYGEFRA 513

Db 499 TVGSSNSVNLDFGDIISAVVYVNTKQRYNSESPLVQMAITLSSIGN-----KSV 552

Db 514 NTIAFTSSGQGGDLPPEKTYKIGDYVWEDVDKGIQNTNDRKPLSNLVLTLYTP 573

Db 553 NALGFTNNGSGAG---QEVYKIGYVWEDVTNKNVQEL--GEKGVNTVTVV-FD 605

Db 574 SKSVR---TDEDGY 585

Db 606 NTKVGEAVTKEDGSY 620

RESULT 5

Db 603

RESULT 6
C89824
hypothetical protein sdrc [imported] - Staphylococcus aureus (strain
C:Species: Staphylococcus aureus

C:Date: 10-May-2001 #sequence revision 10-May-2001 #text change 22-Oct-2001
 C:Accession: C89824
 R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguma, A.; Mizutani-Ul, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiratsugu, K.
 Lancet 357, 1225-1240, 2001
 A:Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.
 A:Reference number: A89758; MUID:21311952; PMID:11418146
 A:Accession: C89824
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1953 <KUR>
 A:Cross-references: GB:BA000018; PID:g13700453; PIDN:BA841750.1; GSPDB:GN00149
 A:Experimental source: strain N315
 C:Genetics:
 A:Gene: sdrc

Query Match 15.9%; Score 490; DB 2; Length 953;
 Best Local Similarity 28.2%; Pred. No. 7.7e-14;
 Matches 162; Conservative 83; Mismatches 236; Indels 94; Gaps 22;

```

42 GLEKSEDETESTNDNEATFLKTPQDNTHLEEEVKESSVSSSIDTAAQPSH 101
| : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : :
45 GHEAKAEHTNGELNOSKNETT---APEENK--TTEKV--DSRLKDNTOATADPKV 96
| : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : :
97 T-----MDSATVKETSSNMOS-----PQNAVASQTQTSNV 129
| : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : :
102 TTINREESVOTSDNEDSHVDFANSKIKESNTSGKEENTTEOPNKVEDSTTSOPSGY 161
| : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : :
162 TNIDEK---ISODELLNPINEYEN-KAPLSTSAQPSIKRYVNOJLA-EQSSNVH 216
| : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : :
130 TNDKSTSTYSNETSKNL--TQAKNVSTTPKTTIKQALNMAVNTAAPOQGTNVND 187
| : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : :
217 LKVTQOST-----TEGYDSEGVKADAEMLDYVEEVDKKSQDTWT 263
| : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : :
188 KAEFTNIDIAIDKGHNKKTGTGTEFWATSDVAK-----LKAVTTIDDSKEDDTT 239
| : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : :
264 VIDDKTVPDLDSTFTPIKIDNCEIATGYDNKNKQITFTDYVKEENIKAHK 323
| : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : :
240 FKYGQFREGSVLPSQTOQLYNAQNIILAKIYDSKTTTFTYTNVYDQVNVGSPFE 299
| : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : :
324 LSYIDSKVPPNNNTKLDVEYKALSSVNTTVEYQRPENKFTALQSFTHIDKNT 383
| : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : :
300 QVAFARENATTKTAYKMEVTLGNDTYSKDIYD---GNQSGQOLISSTNNIMEDLS 356
| : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : :
384 VEGTIIYN-PLRYSAKETNV-NISGNDEGSLIIDSTIKYKVGNDQNLPSNRIDY 441
| : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : :
357 RNMTVYNQPKTYYTETETNLT-----GKFNPAKAKKIYEVDDQNFVDS-FTPD 410
| : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : :
442 SEVEDVTND-DYAQLGNNDVNIINFGN---IDSPYIKVYSKDPKDDYTTIOQTVM 496
| : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : :
411 SKLADVTGQFDVYISNDNKATATVDLNGQSSSKQYIIQVAVPDNSSTNGKIDTLET 470
| : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : :
497 QTINETGEFRASYNTIAFTSSGQGGDLPEKTYIKGYVMDVKKDGIQNTNDN 556
| : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : :
471 QNGKSSWSN-----SYSNVSSSTAND-----QKRYNLGDIWEDTKKQDQ--N 516
| : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : :
557 EKPLSNVLTLYTPDGT-KSVRTDEGKYQFDGV 590
| : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : :
517 EKGIGYVILKDSNGKEDRTTIDENKQIFGL 551
| : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 7
 541539
 fibrinogen-binding protein - *Staphylococcus aureus*
 N:Alternate names: clumping factor
 C:Species: *Staphylococcus aureus*
 C:Date: 13-Jan-1995 #sequence revision 13-Jan-1995 #text change 15-Oct-1999
 C:Accession: S41539; S36630
 R:McDevitt, D.; Francois, P.; Vaudaux, P.; Foster, T.J.
 Mol. Microbiol. 11, 237-248, 1994
 A:Title: Molecular characterization of the clumping factor (fibrinogen receptor) of *Staphylococcus aureus*

A:Reference number: S41539; MUID:94224142
 A:Accession: S41539
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1933 <MCD>
 A:Cross-references: EMBL:Z18852; NID:g397525; PIDN:CAA79304.1; PID:g397526

Query Match 14.6%; Score 452; DB 2; Length 933;
 Best Local Similarity 26.6%; Pred. No. 3.1e-12;
 Matches 154; Conservative 105; Mismatches 234; Indels 86; Gaps 26;

```

8 SDEKNDVNNNOINFDNNQIIKKEETNNYDGEKSEESTESTNDNEATFLK 67
| : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : :
51 SNEKSNDSVSAAPKDDTVN-----SDKITSNTN--NGETSVAQ 91
| : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : :
68 TPQDNTHLEEEVKESSVSSSIDTAAQPSHTTINREESVOTSDNEDSHVDFANS 127
| : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : :
92 NP-----AQGETTQSSSTNATTEETPVGAEATTTTNOANTPATQOS-SNTNAELVN- 143
| : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : :
128 KIKESNTSGKEENTTEOPNKVEDSTTSQPGYINIDEKISODELLN-LPINEYENKA 186
| : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : :
144 --QTSNETFNDTNTV-----SSVNSPQNSTNAENVSTQDTSTEATPSN--NDS 189
| : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : :
187 RLSTTSAPSIKRYTVN-----OLAAEGSNVNLIKYTDQ--SITEGYDSEGV 235
| : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : :
190 APOSTDASKDDVNOAVNTSAPRMRAFLAAVAADPAAGTDITNLTNTVYGI-DSGTT 248
| : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : :
236 IKAHDAENLYDVEEVDKVKSGDTMTVIDKNTVPDLSFTPIKIDNCEIATG 295
| : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : :
249 VYPHQAGYKLVKGFVPSAVKAGDGFKITVRELNLNVTSPAKVFLMAGD-QVLANG 307
| : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : :
296 TYDNKNKQITFTDYVKEENIKAHKLTSTIDSKVPPNNNTK-LDVEYKALSS--VN 352
| : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : :
308 VIDS--DGNVYIYFTDYVKNKDDVAKATLMPAYID---PENAKKGVNVLATGIGSTAN 362
| : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : :
353 KTIYVEYQRPENKFTALQSFTHIDKNTKNTVQTIYINPL--RYSAKETNVNISGNDE 410
| : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : :
363 KTVLVDYKYGKRYNLSTIGTIDQIKNTNTTQTIYVNSGQNVAVLTLGKAKNTBS 422
| : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : :
411 GSTIIDSSTIKYKVGNDQNLPSNRIDYSEVEDVTNDYAQLGNNDVNIINFG--- 466
| : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : :
423 NALIDQNTSIKYYKVNDNAADSESYFV--NPENFEDTVNSVNTFPNPQKYEFVTPDD 481
| : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : :
467 NIDSPYIKVYSKDPN-KDVTYTIQTYMTQTIYEGE--FRASYNTIAFTSSSG 523
| : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : :
482 QITTPYIVVNGHIDNSKGD-----LALRSTLYGYSNIIIRSMWSMDEVAFFNNGS 534
| : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : :
524 OGQG-DLP--PEKTYKIGDY--VMEVDKRD-GIQNTNDN 556
| : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : :
535 SGDGIDKPVYEPQDEPEIEPIEDSDSDPSGSDSGDS 573
| : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 8
 D89852
 fibrinogen-binding protein A, clumping factor [imported] - *Staphylococcus aureus* (str
 C:Species: *Staphylococcus aureus*
 C:Date: 10-May-2001 #sequence revision 10-May-2001 #text change 22-Oct-2001
 C:Accession: D89852
 R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; C.
 ma, A.; Mizutani-Ul, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.
 C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiratsugu, K.
 Lancet 357, 1225-1240, 2001
 A:Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.
 A:Reference number: A89758; MUID:21311952; PMID:11418146
 A:Accession: D89852
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1989 <KUR>
 A:Cross-references: GB:BA000018; PID:g13700678; PIDN:BA841750.1; GSPDB:GN00149
 A:Experimental source: strain N315
 C:Genetics:
 A:Gene: clfa

Query Match	14.6%	Score 452;	DB 2;	Length 989;
Best Local Similarity	25.9%	Pred. No. 3.4e-12;		
Matches 152;	Conservative 109;	Mismatches 224;	Indels 102;	Gaps 27/;

[illegible]

```

RESULT      9
P90070
    jumping factor B [imported] - Staphylococcus aureus (strain N315)
    Species: Staphylococcus aureus
    -Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-oct-2001
    C.Accession: F90070
    R.Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogino,
    M. A.; Mizutani-Uji, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
    C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, I.; Hiramatsu, K.
    Lancel 357, 1225-1240, 2001
    A.Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
    A.Reference number: A89758; MUID:21311952; PMID:11418146
    A.Accession: F90070
    A.Status: preliminary
    A.Molecule type: DNA
    A.Residues: 1-877 <Kur>
    A.Cross-references: GB:BA000018; PTD:g13702568; PIDN:BAH43728.1; GSPDB:GN00149
    A.Experimental source: strain N315
    C.Gene: clfB

```

Query Match	13.1%;	Score 405.5;	DB 2;	Length 877;
Best Local Similarity	24.4%;	Pred. No. 2.8e-10;		
Matches 144;	Conservative 82;	Mismatches 210;	Indels 153;	Gaps 23

Oy 4 HHHHSDSEKEDVNI---NNNOSINTDDNNQIIKKEETNNYGIKRSKSDPRESTTNDEN 60
 Db 39 NHQAOASQSDNDTQSSKNNAASADEKKNM1-----ETQLTNTANDNSDIANTNSANVND 94
 Oy 61 EATFLQKTPDNDTHLTREEEVKSSSVSSSSIDTAQOPSHPTINREESVQTSNDVEDSH 120
 Db 95 STTKPMSTQTSNTTTE-----PASTNETPQPTAIKNQ---ATAAKMODOT 137
 Oy 121 VSDPANSIKESNTESGEENTIEDQPNKVKEDSTTSQSGTINDEKISNODELLNLPIN 180
 Db 138 VPOEANSQVDRKKT---NDANSIATNSSELKNSQT-----LDLP-- 172
 Oy 181 EYENKARPLSTTSAQ---PSIKRYTVNQLAAEQ-----GSNVHLIKVTDQSITE 227
 Db 173 ----QSSPQTSINMOGTSKPSVPRPAVASLVAEYVVAADAAGKTNN-----DKVTAS 222
 Oy 228 GYDSEGVIAKHADEANLIYDTFEYDVKVSGDVTV-----DID---KNTVPS 273
 Db 223 NFKLEKTFDPDQSGNTFMANAFVTVDKKGSDYFTAKLPDLNGNGVDYDVSNSNMTNP 281
 Oy 274 DLTDFTLPKIKDNGSELIATGTVDNKNKQLTYPETDQVDKXENKALKLTSTYDASKV 333
 Db 282 ----TADIKSTNGDVAAKATFDILTRTYFTVDNKNKENINGOFSLPLETRAKA 334
 Oy 334 PNNNTKLDVEKXTALSSVNTKITVEYQRP---NENRTANLQSMETNTDTKN--HTVEQT 387
 Db 335 PKSGP-YDANINIADEMENNKITVYSSPIAGIDKPGANISQIIGDVASGQWTYKOT 393
 Oy 368 IYINLPKRSKAKETNNVNSGNDEGSTIIDDS-----TIKYKVQVDNQLNLPDS---- 435
 Db 394 VFVNPKQRYALGNTWYLIKGYQDK-----IEESSGKVASATDTKLRIFEVNDISKLSDYAD 449
 Oy 436 -----NRIDYSEYEDVTNDDDYAQLGNNDVAINFEGNIDSPYIIKIKSKYD 481
 Db 450 PNDLSLKEVTDFQFKRRI-----YE-----HPVVASIKEGDITKRYVVLVEGRYD 494
 Oy 482 PNKDDYTT--IQQVYTWQTTINETYGEFRPTASDYNTIAFTSSGGQGD 528
 Db 495 NTGKMILKQVIOENDVDPYT-----NRDYSIRGMNNENYVRGGGSADGD 538
 RESULT 10
 S19702
 fibronectin-binding protein B - *Staphylococcus aureus*
 C:Species: *Staphylococcus aureus*
 C:date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 15-Oct-1999
 C:Accession: S19702
 R:Joensen, K.; Signaes, C.; Mueller, H.P.; Lindberg, M.
 Eur. J. Biochem. 202, 1041-1048, 1991
 A:Title: Two different genes encode fibronectin binding proteins in *Staphylococcus aureus*
 A:Reference number: S19702; MUID:92111475
 A:Accession: S19702
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-940 <JOB>
 A:Cross-references: EMBL:X62992; NID:g49040; PIDN:CA44726.1; PID:g581562
 C:Keywords: fibronectin binding

Query Match	12.28;	Score 376;	DB 2;	Length 940;
Best Local Similarity	23.48;	Pred No. 5.6e-09;		
Matches 138;	Conservative 217;	Mismatches 217;	Indels 124;	Gaps 22

[illegible]

Db 136 E-----IKPRMRSKSTDTAVAEK 153
Qy 211 -----GSNNHLIKYTDOSITBEGYDSEVIKAHDAENLIYDTEFVDKVSQGT 261
Db 154 EYEEETKATCTDVTKNVEEAGESEIYGHKODITVNVPHNAEFYTLKKMKFEGIKAGDY 213
Qy 262 MIVDIDKNTVPSDLTDSFTIPKIKDNGSEIATGYDNKKKCLTPTFDYDVENIKAH 321
Db 214 FDTTLDSDNETHGISITLRKVPETKSTDGQVMAAGEIIGENK-VRYTFKEVQEKDULAE 272
Qy 322 LKLTYSIDKSKVPNNNTKLDVEYKTALESYKNTITVEY---QRPNNRATLMSMTNID 378
Db 273 LSLNLFIDPTTYQKGNQ-NVEYKLGSETVSKIFENIQYLGVRDMNGVLAN-----GRID 326
Qy 379 TKMHT---VEQTIYINPLKSAKETVNVISGNDEGTTIIDSTIIKYKRVGNOLPDS 435
Db 327 TLNRKVDGKFSHFAVYKMRNNOSL--SSVTVIGQVTKGKRKPGVNPYKVVYHHISDLSAES 384
Qy 436 --NRIYDSEYEDVNDYDQAOLGNNDNINFGNIDSP--YIIKIVISKYDPNKDDYTTIQ 491
Db 385 VYAKLDDVSKFEEDVTDMMSLDPTNGSYSLNFNNLDOSKRYVYIKYEGYDMSASN----- 439
Qy 492 QVVTMOTITNEYTGEFRTA--SYDNTIAFSTSSGQGG-----DLPEKTYKIG 538
Db 440 --LEFQTHLFGYNYNYTSLTMKNGVAFYSNNAOGDGKDKLKEPIIEHSTPIELEPKSE 497
Qy 539 DYWMEDVDKGIQNTDNDNEKPLSNVLTFL-----YPDGTSKVRDDED 582
Db 498 PVEKHEHLLGTIEESNDS-KPIDFEXHTAVEGAEGHAEGT---IETED 542

RESULT 11

G90053
Hypothetical protein fnbB [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #ext_change 22-Oct-2001
C:Accession: G90053
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc ma, A.; Mizutani-U, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: G90053
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1961 <KUR>
A:Cross-references: GB:BA000018; PID:g13702452; PIDN:BA843593.1; GSPDB:GN00149
A:Experimental source: strain N315
A:Genetics:
A:Gene: fnbB

Query Match 11.2%; Score 345.5; DB 2; Length 961;
Best Local Similarity 22.2%; Pred. No. 1.2e-07;
Matches 125; Conservative 132; Mismatches 219; Indels 87; Gaps 24;
Qy 81 KESSVSSNSIDTAQPSHTTINRESVOTSDNVEDSHVSDANSKIKESNTESGKEE 140
Db 33 KEAASQNNNTYVESGS-SATESKASETOTTNNV-----NTIDETOSYSA--- 78
Qy 141 NTIEOPKRVKEDSTTSQPSGY--TNIDEKISNODELINLPINEYENK-----ARPLST 192
Db 79 TSMOPKSKQVYTEAPTVQAPKVFTEKMSQEDLPSEKRVADDETTGTQVDAQPSNV 138
Qy 193 SAQPSIR-----VYVNOJLAE--OGSNVNLIKYTDOSTIRGYDSEGVYKAHDAEN 243
Db 139 EIKPRMRSADVAVSEKEVAEEKATGTDVTKNVEYTESSL-HGHKSDSNIVAPHNAQR 197
Qy 244 LIYDVEEDVDKVSQGTMTVVDIDKNTVPSDLTDSFTIPKIKONS-GEIATGYDNKK 302
Db 198 VTLKMKFEGIKAGDYFDTTLDSDNETHGISITLRKVPETKSTEDKVMANGOVINE-R 256

Qy 303 QITTFEDVDKVENIKAHKLKLSYIDKSKVPNNNTKLDVEYKTALESYKNTITVEYQRP 362
Db 257 TIRYFTDYINNKKDLAEINLNLFTDPTTVKQSSQ-KVEVTLGQKVSKEPDYTLDS 315
Qy 363 NENR---TNLMSMTNIDTKNH--TVEQTIYNPLKSAKETVNVISGNDEGTTIID 416
Db 316 VKRMGVTVN-----GRIDLKEEGKFSHFAVYKMRNNOSL--TSVTVIGQVYSGVKOSA 368
Qy 417 DSTIITVYKVGDNOLPDS--NRIYDSEYEDVNDYDQAOLGNNDNINFGNIDS--PY 472
Db 369 NNPTVYKHHIGSDELAESYAKLDTSKEPDYEVNLSYSTNGGYTLMLGDLNDSKY 428
Qy 473 IIVISKYDPNKDDYTTIQVVTMOTITNEYTGEFRTASV-----DNTIAFSTSS 522
Db 429 VIKYGEYIDQNAKD-----LNFRTHLSGYHAKYIPYIPYIPYIPYQVLTNNKVAFTSN 481
Qy 523 GQGGQ-DLPEKTYKIGDYWMEDVDK-----GIQNTDNEKPLS-NVLTFLTY 570
Db 482 AKGDGKRPNDPIEKSEPIDLIKSEPPVEKHEHLLGTIEESNDS-KPIDFEXHTAVEGA 540
Qy 571 DGTSKSVRDEDEGKYQPDQVQD 593
Db 541 EGHAEGLIETEE-----DSIHVD 558

RESULT 12

H90053
Hypothetical protein fnbB [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #ext_change 22-Oct-2001
C:Accession: H90053
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Cui, A.; Mizutani-U, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: H90053
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1038 <KUR>
A:Cross-references: GB:BA000018; PID:g13702453; PIDN:BA843594.1; GSPDB:GN00149
A:Experimental source: strain N315
A:Genetics:
A:Gene: fnbB

Query Match 10.6%; Score 326; DB 2; Length 1038;
Best Local Similarity 23.1%; Pred. No. 8.8e-07;
Matches 139; Conservative 102; Mismatches 215; Indels 146; Gaps 30;
Qy 42 GIKRSEDRTSTNTDNDENATFLQKTPQDN-----THLTPEYKESSEVSSNSIDTAQ 97
Db 30 GQDKKAASEQKTYEEN-----GNSATDNKTSETOTTNNHIEETOSYNATV--TE 82
Qy 98 QPSHTT-INRESVOTSDNVEDSHVSDANSKIKESNTESGKEENTIEOPNKVEDSTTS 156
Db 83 QPBNATQVTEENAPKAVQ-----APQTAQPAVVEYKKE--EK-QVKE--TT 125
Qy 157 QPSGYTNIDEKISNODELINLPINEYENK-----ARPLSTSAQPSIKRYTVNOJLAE 208
Db 126 QP-----QNSGNQNRQVDTLPKVVQNOQGTETQVEVAQPRTASESKRPVTR-SADVAEA 178
Qy 209 EOGSNVNLIKYVDOSTIEGYDSEGVYKA-----HAENLIYVTEFVDKVSQGT 260
Db 179 KEASDVSE-VKGD--VTSKTVYESGSEIAPQGNKVEPHAGQRYVLTLYKKFADGLKRGD 235
Qy 261 TKMHTVDIDKNTVPSDLTDSFTIPKIKONS-GEIATGYDNKKQITPTFDYDVKY 315
Db 236 YPFTLSNNVNTYGVSTARKVPEIKNSVYMAAGEILGN-----NIRYTTNNELEHK 288
Qy 316 ENIKAHKLKLSYIDKSKVPNN-----NTKLDVEYKTALESYKNTITVEYQRPENRTANL 370

Db 289 VEVTANLEINLEIDPKTVQSNQGEOKITSLNGE-----ETEKTIPYVNPVNGVNSYTNV 342
QY 371 QSMETNIDTNRHVEQTIYINPLAYSARENVNINSGNDESTIITDSTIIKYK-VGDN 429
Db 343 NGSIETPKSNKFTTHAIYIKPMNGOSMT-VSVTGLTESSNLAGOPVTKVYEVLEKK 401
QY 430 ONLPDSNRIT---DYSEEDVTNDYAOALG--NNNDVINFGNIDSPYIITKISKYDPN 483
Db 402 DELPQS--VVAISDITKFKDYTKEMANGKLSVQNGSYSLMDKIDTYVHYHGEYILOG 459
QY 484 KDPYTTIOQTVMQTTINEYTGFEPTASDYNTIAFSTSSGOGDLPPEKTYIKIGDYWE 543
Db 460 SDQY-----NRTELYGY-----PERAYK-SYVYV- 483
QY 544 DVKDGIONTNDNEKPLSNVLTLYTPDGTSKS--VRTDEGKY-----QPDGV 590
Db 484 ---GGYRLTWDN-----GLVLYSNKADGNKNGOIIDNDFEYKEDTAKTMSGOYDAK 534
QY 591 QY 592
Db 535 QY 536

RESULT 13
A32192
fibronectin-binding protein - Staphylococcus aureus
C:Species: Staphylococcus aureus
C:Date: 31-Jul-1989 #sequence_revision 31-Jul-1989 #text_change 16-Feb-1997
C:Accession: A32192
R:Signes, C.; Raucet, G.; Joensen, K.; Lindgren, P.E.; Anantharamiah, G.M.; Hoeck, M.
Proc. Natl. Acad. Sci. U.S.A. 86: 699-703, 1989
A:Title: Nucleotide sequence of the gene for a fibronectin-binding protein from Staphylococcus aureus
A:Reference number: A32192; MUID:8908998
A:Accession: A32192
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1018 <SIG>
A:Cross-references: GB:J04151
C:Keywords: fibronectin binding

Query Match 9.8%; Score 304; DB 2; Length 1018;
Best Local Similarity 21.2%; Pred. No. 7.5e-06;
Matches 12; Conservative 12; Mismatches 24; Indels 104; Gaps 26;
QY 42 GIERSEDRSTSTNVDENEATFLQKTPDN---THLLEEVEKSSSVESNSSIDTAQ 97
Db 30 GQDEMAASEQKTTVEEN---GNSATDKTSETOTATNVNIIETQSYNATV--TE 82
QY 98 QPSHTT-INREE---SVQTSNVEDSHVDFPANSKIKESNTESGKEENTIQPNKVKEDS 153
Db 83 QPSNATQVTEEAARKAVAPOTAPANIEYKEEYKE-----EAKPKVKEET 130
QY 154 TTSQSG---TNTIDKISNDELINLPINEENKARLSTSAQESIKRYT-----V 203
Db 131 OSQNSGGQROVDLPKAKATNOVAETOV---EVAQPTASSESKRYTRSDAFAEKA 186
QY 204 NQLAEOGSNNVHLIKYDOSITEGVDSEGIKAHDAENLYDTFVEDDKVKSQDPT 263
Db 187 SNAKETGTDTYSKATVEIGSI-BGHNNKV-EPHAGQRAVLAKKLPENLHGGDVF 244
QY 264 VDIKNTYPSDLTDSFTIPKIKDNS---GEIATGYDNKKNKQITTFDDYVDKYE 318
Db 245 FTLSNNVTGHTARKVPEIKNGSVNATGEVLEGG-----KIRYTFNDIDKVDV 297
QY 319 KAHLKLTGYIDSKSKVPNNNTKLDVETKALSSVN-----KTTVEYCPNENRNTANLOSM 373
Db 298 TAELEINLFIDPKTYQTGN-----QTTITLNEQTSKELDVKYKQIGIYVYANLNGS 351
QY 374 FTNIDTKNHTVEQTIYINPLAYSARENVNINSGNDESTIITDSTIIKYK-VGDN 432
Db 352 IETFKANNRSHVAFIKP--NNGKTSVTYVGTLMKSGNSQNGQPKVIRIFEYLCNNEDI 409

QY 433 PDSNRIT---DYSEEDVTNDYAOALG--NNNDVINFGNIDSPYIITKISKYDPNKD- 485
Db 410 AKS--VVAISDITKFKDYTKEMANGKLSVQNGSYSLMDKIDTYVHYHGEYILOG 467
QY 486 -DYTTIOQTVMQTTINEYTGFEPTASDYNTIAFSTSSGOGDLPPEKTYIKIGDYWE 544
Db 468 VDFET--QWVGHPEDLYKYRYDGYTLWMDNGLVLYSNKANGN----- 508
QY 545 VDKG--IONTNDNEKPLSNVLTLYTPDGTSKSVRDEDEKQYQPDGVQVD 593
Db 509 -EKNGPIION-NKFEYKEDTIKERTLT--GOYDKMLVYTFEEYDSSTLID 555

RESULT 14
T30856
protein F2 - Streptococcus pyogenes
C:Species: Streptococcus pyogenes
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 11-May-2000
C:Accession: T30856
R:Jaaffe, J.; Natanson-Yaron, S.; Caparon, M.G.; Hanski, E.
Mol. Microbiol. 21, 373-384, 1996
A:Title: Protein F2, a novel fibronectin-binding protein from Streptococcus pyogenes,
A:Reference number: 220907; MUID:97011581
A:Accession: T30856
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1039 <AAF>
A:Cross-references: EMBL:U031980; NID:91654115; PID:91654116; PIDN:AA044522.1
C:Genetics:
A:Note: PRRT2

Query Match 7.7%; Score 239; DB 2; Length 1039;
Best Local Similarity 22.6%; Pred. No. 0.0046;
Matches 93; Conservative 79; Mismatches 168; Indels 72; Gaps 18;
QY 224 SITEGYDSEGVIRAHND---AENLIYDTEFEVDKVS-----GDTVMYDIDKNTVPSDL 275
Db 38 SVSKRYGEQERTSNADRYNNHAAVFKMSFELKOKDKSETINPGDTEVYQLDRRLNPKGI 97
QY 276 TDSFTIPK-INDNSGELIATGYDNKKNKQITTFDDYDKYENIKAHKLTYSIDSKVP 334
Db 98 SOD--IRKIIYDSENSPLAIGKDAKTHQLYTFTNYIAGIDKQVLSAELFLFNKREVL 155
QY 335 NNNKRLDVEYKTAALSVKKTITVEYQRPNE-----NRTANLOSMFTNIDKRNHVE 385
Db 156 ENFNISDRKSTIGQETIYKGTAVNLVNGNESTKESNYITNGLSNVGSGISEYNETGEFV 215
QY 386 QTIYINPLAYSARENVNINSG-----NGDEGSTIITDST--TIKYKYVGNQNPDSNR 437
Db 216 WYVYVNPRTNIPYAVLMLMGFAKRTAGQENDNSLVSAOGLGYDIEVPHNYRLPTSYG 275
QY 438 IYDSEYEDVTNDYAOAL-----GNNDVININFG-NID-SPYIITKISKYDPNKDVTY 489
Db 276 V-DISRL-NLKLDLAKLPQSGTQGANRKLRIIDRENIQGAFAVYKVKAKDQSKR--- 330
QY 490 IQQVYTMQTTINEYTGFEPTASDYNTIAFSTSSGOGG-----DLPEKT 534
Db 331 ---LIVQSHLSFNNMGSYKTLRPNSHVSTNEIALSPKSGSGSGTSEFTKPSITVANL 386
QY 535 KYIGDYVEDVDKDI-----QNTNDNEKPL--SNVLTLYTPDGTS 574
Db 387 KRVAOLRFKKVSTDNVPLPEAAFEILRSSNGSOKLEASSNTOGETHFRDLTS 438

RESULT 15
T30290
AAS surface protein - Staphylococcus saprophyticus
C:Species: Staphylococcus saprophyticus
C:Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 02-Sep-2000
C:Accession: T30290
R:Hell, W.; Meyer, H.G.W.; Gattermann, S.G.
Mol. Microbiol. 29, 871-881, 1998

A:Title: Cloning of aas, a gene encoding a Staphylococcus saprophyticus surface protein
A:Reference number: Z20809; MUID:96389318
A:Accession: T30290
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1463 <HELD>
A:Cross-references: EMBL:AJ000007; NID:e1295629; PID:e1295630; PIDN:CAA03852.1
C:Genetics:
A:Gene: aas

Query Match: 7.3%; Score 225; DB 2; Length 1463;
Best Local Similarity 20.9%; Pred. No. 0.029;
Matches 143; Conservative 105; Mismatches 263; Indels 172; Gaps 33;

```
OY      8 SDEKNDVNNNOI-----NTDNNQIIRKEETNNYDGIKRESDRTSTST--- 55
DB      180 NTDENSDVANNQNEPVAQNDKAETSNEDVASSDVAKQDGTSDNADSDLTQDNESETLND 239
        56 -NYDNEATFLQKTPODNH-----LTE--EEVKSSSVSSN---SIDTAAQPSH 101
DB      240 NAVSSNEDVASSDVAKQDGTSDNADSDLTQDNEPVAQNDKAETSNEDVASSDVAKQDDTH 299
OY      102 TTN-----REESVOTSDNVEDSHVDFANKIKESNTESG-----KEENTIEOPN- 147
DB      300 SPANASDVADQNESETQNDNMESSNEDVASSDVAKQDDTHSDANASDVAKQDNESETQND 359
OY      148 -----KVKEDSTS--QPSGYTNIDEKISNODELNLPINEYENKAPLSTT 192
DB      360 KAETSNEDDVASSDVAKQDDTHSDANASDIADQNESVQAQDKA-----ETSNEDVASS 411
OY      193 SAQPSIKRTVNOQLAEQGSNNHLIKYT--DQSITEGYDS---EGYIK-AHDAENLI 245
DB      412 DKQDDTHSDANASDIADQNESATQDDKATSKEDDVASNDKQDNKAVSNIIKEASTAENKV 471
OY      246 YDVT--EYVDKVKSGDFK--TVDIDKNVPSDLTDSFTPIKIDNSGEIITATGYDNKN 301
DB      472 QPRTSAKYTPKLRVATTSANTAVATRSAYTKAATTRALPKY-----S 515
OY      302 KOITTFDYDYDKYENIKA--HLKLTYSIDSKVPPNNNTKLDVEYKTALSSVNKTIVE 358
DB      516 PKVNSSINNYIRK--NNFKAPNEYQDIANYLPQYNRYGKPEGIYMHDTA--NDNSTITGE 572
OY      359 YQRPNENRTANQSMFTNIDTKNHTVEQRTIYNPLRYSAKETNWNISGNGDEGSTIIDS 418
DB      573 INYMKNNYT---SAFVHAHYVDGDRILETANTDYLAWGA-----GQANDRF 615
        419 TIKKYKVGDNQNLDPDSNRIYDYSEYEDVTNDYQAQLGNNNDVNTNFGNIDSPYIIKYVS 478
DB      616 IHVELVHTHDYDFARS--INNYADYA--ATNQLYGLVDP---SAEYDGVGTWTHQAVS 669
OY      479 KY-----DPN-----KDYTTIQQVTVMQTTINEYTGERTASYDNTIAFST- 520
DB      670 NYLGSGSDHSDPHGYLAHNYSYDELYDLIEKYLITKG-----QAAAGTSSGSGTG 721
OY      521 ----SSGQGO--GDLRPEK--TYKIGDYWEDVDKDGIONTNDNEKPLSNVLTITLYPDGT 573
DB      722 GTGSGTSGSNTGTTTPSKSGYKVV-----TENNGVGRINSK---NDGLTYTVYDQNG 770
OY      574 SKSVRTDED-----GKYOF 587
DB      771 KTTDRVNOQLTKYTKSATLQKEQY 793
```

Search completed: July 30, 2002, 10:31:29
Job time: 154 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 30, 2002, 10:31:56 ; Search time 28.43 Seconds

(without alignments)
807.622 Million cell updates/sec

Title: US-09-147-405-11

Perfect score: 3087
Sequence: 1 HHHHHHPSDEEDKNDVNNNN.....SKSVPRDEQKQFGDGVQVD 593

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwisProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES:

Result No.	Score	Query	Match Length	ID	Description
1	304	9.8	1018	1	FNBA_STAU
2	214	6.9	1093	1	PI4K_DICD1
3	203	6.6	537	1	ARP_PLAFA
4	199.5	6.5	1139	1	HMW1_MYGE
5	197	6.4	918	1	YMB_CAEEL
6	195	6.3	731	1	BAF1_YEAST
7	192.5	6.2	578	1	LIPA_MYCPU
8	189.5	6.1	1664	1	INT1_CANAL
9	188.5	6.1	1435	1	EBAL_PLARC
10	186.5	6.0	3110	1	LM42_HUMAN
11	186	6.0	1381	1	YB7_YEAST
12	185.5	6.0	1165	1	YNF4_YEAST
13	184.5	6.0	1744	1	TANA_XENLA
14	184	6.0	817	1	YGA4_YEAST
15	184	6.0	1251	1	YMB2_PLAVB
16	184	6.0	1658	1	YMB7_YEAST
17	183.5	5.9	2009	1	SEC7_YEAST
18	183.5	5.9	2452	1	RPL1_PLAFD
19	182.5	5.9	853	1	YCG1_YEAST
20	182.5	5.9	1875	1	MLP1_YEAST
21	182	5.9	1419	1	ALAL_CANAL
22	182	5.9	1460	1	NI59_YEAST
23	181.5	5.9	1093	1	KD5_CAEEL
24	180	5.8	1253	1	DSP_HUMAN
25	178	5.8	960	1	YMX6_YEAST
26	177.5	5.7	1196	1	BXCN_GLOBO
27	176.5	5.7	1637	1	MRSP_STAU
28	176.5	5.7	1803	1	YUJ3_YEAST
29	175.5	5.7	1183	1	CNA_STAU
30	175	5.7	1147	1	CGA1_HELPY
31	174.5	5.6	682	1	NI5P_LACLA
32	174	5.6	1466	1	SPA2_YEAST
33	174	5.6	2022	1	ANL1_ONCVO

34	173.5	5.6	967	1	Y5G0_CLOAB	P33747 clostridium
35	173.5	5.6	1260	1	AL51_CANAL	P46590 candida alb
36	173	5.6	827	1	CSG_HALVO	P25062 halobacteri
37	172.5	5.6	688	1	LIP_STARP	P02510 staphylococ
38	170.5	5.5	970	1	PSU1_YEAST	P33550 saccharomyc
39	170.5	5.5	1435	1	LTEL_YEAST	P07866 saccharomyc
40	170.5	5.5	1636	1	BUD3_YEAST	P25558 saccharomyc
41	170	5.5	2334	1	WAPA_BACSU	P07833 bacillus su
42	169.5	5.5	571	1	TACY_STRPY	P21131 streptococc
43	169.5	5.5	1020	1	BCA_STRAG	P002197 streptococ
44	169.5	5.5	2869	1	RBP1_PLAVB	P00798 plasmodium
45	169	5.5	954	1	BRL1_YEAST	P47134 saccharomyc

ALIGNMENTS

RESULT 1
ID FNBA_STAU STANDARD; PRT; 1018 AA.
AC PI4738:
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Fibronectin-binding protein precursor (FNBP).
GN FNBA.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCYC 8325-4;
RX MEDLINE=89098998; PubMed=2521391;
RA Signaes C., Raucel G., Joensson K., Lindgren P.-E.,
RA Anantharamiah G.M., Hoeoek M., Lindberg M.;
RT "Nucleotide sequence of the gene for a fibronectin-binding protein
from Staphylococcus aureus: use of this peptide sequence in the
synthesis of biologically active peptides".
RT Proc. Natl. Acad. Sci. U.S.A. 86:699-703(1989).
RL -1- FUNCTION: THE ABILITY OF BACTERIA TO BIND FIBRONECTIN HAS BEEN
PROPOSED AS A VIRULENCE FACTOR ENABLING BACTERIA TO COLONIZE
WOUND TISSUES AND BLOOD CLOTS. BINDING OF PLASMA FIBRONECTIN TO
THE BACTERIAL SURFACE MIGHT BLOCK ADHESION RECEPTORS ON S.AUREUS,
THIS REPRESENTING AN IMPORTANT DEFENSE MECHANISM AGAINST TISSUE
INVASION.
CC CC
CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Cell wall.
CC -1- SIMILARITY: TO OTHER STREPTOCOCCAL AND STAPHYLOCOCCAL PROTEINS
CC IN THE REGION OF THE MEMBRANE ANCHOR.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: J04151; AAA26632.1; -
CC InterPro: IPR004237; Fn_bind.
CC InterPro: IPR001899; Gram_pos_anchor.
DR Pfam: PF00746; Gram_pos_anchor. 1.
DR Pfam: PF02986; Fn_bind; 1.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; 1.
KW Signal; Repeat; Cell wall; Transmembrane.
FT CHAIN 1 1018 FIBRONECTIN-BINDING PROTEIN.
FT SIGNAL 37 993 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 994 1009 MEMBRANE ANCHOR.
FT TRANSMEM 1010 1018 CYTOPLASMIC (POTENTIAL).
FT REPEAT 545 574 B-1.
FT REPEAT 575 604 B-2.
FT DOMAIN 745 878 4 X APPROXIMATE TANDEM REPEATS,

```

FT REPEAT 745 782 FIBRONECTIN-BINDING DOMAIN.
FT REPEAT 783 820 D-1.
FT REPEAT 821 859 D-2.
FT REPEAT 860 878 D-3.
FT REPEAT 879 948 D-4 (INCOMPLETE).
FT REPEAT 879 948 5 X TANDEM REPEATS, PRO-RICH (WR).
FT REPEAT 892 892 WR1.
FT REPEAT 893 906 WR2.
FT REPEAT 907 920 WR3.
FT REPEAT 921 934 WR4.
FT REPEAT 935 948 WR5.
FT DOMAIN 982 987 CONSERVED IN GRAM-POSITIVE COCCI SURFACE
PROTEINS.
SQ SEQUENCE 1018 AA; 111780 MW; 58175E002081F1F CRC64;

Query Match 9.8%; Score 304; DB 1; Length 1018;
Best Local Similarity 21.2%; Pred. No. 4,7e-06;
Matches 125; Conservative 121; Mismatches 241; Indels 104; Gaps 26;

42 GLEKSEDETESTNDENAEFLQKTPDND---THLVEEVKSSSVSSSSSITDRAQ 97
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
30 GDDKKAASEQKTTVEEN---GNSANDNKTSETOTATVNHIEQSYNATV--TE 82
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
98 OPSHTT-IRRE---SVOTSDNVEDSHVDFANSKIKESNTESKRENTIEQPNKVEDS 153
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
83 QDSNATQVTEEPKRAVQAPQAPANITVKEVYKE-----EAKPOVKETT 130
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
154 TTSQPSG---VTNIDEKISNOELLNLPINEXENKARPLSTSAOPSIKRVT-----V 203
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
131 OSQDNGSGDQROYDLFPKRAKTONOVAETQV---EVAQPTASESKRPVRSADVAEAKA 186
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
204 NOLAAEGGSNVNHLIKVTQOSTEGYDDESEVYKAHDAENLIYVTFEVDKVKKSDITMT 263
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
187 SNAKVEGTGTVDSKYVVEIGSI-EGHNNTNKV-EPPHAGRAVLKYLKEENGLHOGDYFD 244
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
264 VIDKNTVPSDLDSEFTPIKIDNS-----GELIATGYDNKNKOTTYFTYVDKXENI 318
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
245 FTLSNVNVNTHGVTAKVVEIKNGSVYMATGEVLEG-----KRTYFTNDIEKXVY 297
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
319 KAHKLITSTIDSKVPPNNNTKLDVEYKIALSVN-----KITVEYQRPENRNTANLQSM 373
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
298 TAELEINLEFDPRVQOTNGN-----QITSTLENEQTSKELDVKKDGIGVYANLNGS 351
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
374 FTNIDKNTVEQTIYINLAKYSAKETNNVNSGDEGSIIDDSIIIVYK-VGDNOLV 432
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
352 IETFNKANNRFSHVAFIKP--NNGKTTSVYVGTLMKGSNONGNPKVRIFFELGNNEDI 409
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
433 PDSNRIV---DYSEYEDVTNDYAOLG--NNNDVININGNDSPYIIVISKYDKNK- 485
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
410 AKS--VYANTTDTSKREKETSNGSLNQNNGSYSLNTEINLKDITYVYHIDGETYLANGTDE 467
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
486 -DYTTIQQTVTQTTINEYGEFRFASVDNTIAFSTSSGOGGDUPEPKTYKIGDYVMD 544
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
468 VEFRT-QMVGHPQLKYYRQYGLTMDNGVLVYENKANGN----- 508
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
545 VKDGS--IONTNDNEKPLSNVLVLTLYPQGTSGSVRTDDDKYQFQGYVD 593
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
509 -EKNGFIION-NKEFYKEDTIKETLT--GOYDKNLVTVEEEDYSTLID 555
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 2
PI4K.DICDI STANDARD: PRT; 1093 AA.
AC P54677;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE phosphatidylinositol 4-kinase (EC 2.7.1.67) (PI4-kinase) (PtdIns-4-
kinase) (PI4K-alpha).
GN P1KD OR PI4K.
OS Dictyostellium discoideum (slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostellium.

```

```

OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX3;
RX MEDLINE=96009592; PubMed=7565716;
RA Zhou K., Takegawa K., Emir S.D., Firtel R.A.;
RT "A phosphatidylinositol (PI) kinase gene family in Dictyostellium
discoidum: biological roles of putative mammalian p110 and yeast
Vps34p PI 3-kinase homologs during growth and development.";
RL Mol. Cell. Biol. 15:5645-5656(1995).
CC -I- FUNCTION: ACTS ON PHOSPHATIDYLINOSITOL (PI) IN THE FIRST
COMMITTED STEP IN THE PRODUCTION OF THE SECOND MESSENGER
INOSITOL-1,4,5--TRISPHOSPHATE.
CC -I- CATALYTIC ACTIVITY: ATP + 1-phosphatidy1-D-myo-inositol - ADP +
1-phosphatidy1-D-myo-inositol 4-phosphate.
CC -I- SIMILARITY: BELONGS TO THE PI3/PI4-KINASES FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL: U23479; AAA85725.1; -.
DR DictyDB: DD01102; PI4K.
DR InterPro: IPR00403; PI3_P14_Kinase.
DR Pfam: PF00454; PI3_P14_Kinase; 1.
DR SMART: SM00146; PI3K; 1.
DR PROSITE: PS00915; PI3_4_KINASE_1; 1.
DR PROSITE: PS00916; PI3_4_KINASE_2; 1.
DR PROSITE: PS50290; PI3_4_KINASE_3; 1.
KM transferase; kinase; Multigene family.
FT DOMAIN 146 156 ASP-RICH.
FT DOMAIN 184 200 POLY-ASN.
FT DOMAIN 203 206 POLY-ASN.
FT DOMAIN 227 233 POLY-ASN.
FT DOMAIN 277 315 POLY-ASN.
FT DOMAIN 441 445 POLY-ASP.
FT DOMAIN 455 466 POLY-THR.
FT DOMAIN 494 501 POLY-GLY.
FT DOMAIN 690 696 POLY-THR.
FT DOMAIN 751 754 POLY-GLN.
FT DOMAIN 761 772 POLY-THR.
FT DOMAIN 775 785 POLY-THR.
FT DOMAIN 833 1093 PI3K/PI4K.
SQ SEQUENCE 1093 AA; 123017 MW; 8762BC78355A635 CRC64;

Query Match 6.9%; Score 214; DB 1; Length 1093;
Best Local Similarity 20.2%; Pred. No. 0.028;
Matches 121; Conservative 101; Mismatches 224; Indels 152; Gaps 26;

1 HHHH--HHPSSDEK-----NDVINNOSINTD--DNNOIIRKETNNYDGIERSBD 49
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
167 HHYEDNDHTEENPKKDINSNNNNNNNNNNNNNNNDNNNEILPENSNSNSINDENNOVG 226
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
50 RTESTTND-ENEATFLQKTPDNDTHLVEEVKSSSVSSSSSI--DTAQP----- 99
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
227 NSNNNNNIGENNNIKIDINSON-----KSDNIEITLSTLCEEFKTSPIKDMEN 277
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
100 ----SHITINREESVOTSQNVEDSHVSDPANSKIKESN-----TESGREENT 143
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
278 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 334
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
144 EOPNKVEDSTTSQPSG---VTNIDEKISNOELLNLPINEXENKARPLSTSAOPSIKR 200
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
335 SQPNDEIENITILKRNRIYKVEKKELARLRKFCMSYITGSRPLITPRRSLPS 394
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
201 VTNOLAAEGGSNVNHLIKVTQOSTEGYDDESEVYKAHDAENLIYVTFEVDKVKKSD 260
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
395 PLVYNSGRIQGNVYHILSPSSVDSSTLSIEDDKIIEKEEDNVYVD---DDDEVENSD 451
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```


OY	261	--TMMVVDIDKNNTPVEDLDSFPIPRXONSG--EIIATGTADNKNKIOTFTFDYV	312
Db	452	FIPATTTATTTTTTTTTHPHLSKTSTTGVCINSNSTPPIPINSAGACAGGGGELNH----	506
OY	313	KRYENIKAKHLTSTYDYSKRPVNNNTKL--DVEXKTA-----LSSVNKTIVE	358
Db	507	-GYDI-----SYLDCKCTPPAESKTSJDHDFEELSXHRCDYLNDLFIFDKLAHS	558
OY	359	---YORPENRNTANLOSKEFNIDTKNHTVEQTIYINPLIRYAKETMWNISGNGDGSTII	415
Db	559	KILPFIPIIDLROAKLKH--EISLINIMLPGLYV-PLMOSS-----	596
OY	416	DDSTIIXKYKKGDNDGNPDNSRIYDSE-----YEDTNDYACLGNNDNVINFGNID	469
Db	557	NHHCVNR-----PEEKYKILNSRRVRPFLLVLEYIESEHEHLSN-----	637
OY	470	SPIIIVISKYDPNKDYTTIOQTVMOTINETNYGE-----FRFASDYNTAFSTSS	522
Db	638	---IREVVSSY----LQYTTGSALKKDIDKRKYSEKFSSFLNSSINSTISNSDS	688

SULT			
ID	ARP_PLAFA	STANDARD:	PRT; 537 AA.
AC	P04931;		
DT	13-AUG-1987 (Rel. 05, Created)		
DT	01-NOV-1995 (Rel. 32, Last sequence update)		
DT	01-NOV-1995 (Rel. 32, Last annotation update)		
DE	Asparagine-rich protein (Ag319) (ARP) (Fragment).		
OC	Plasmodium falciparum.		
CC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.		
OX	NCBI_TaxID=5833;		
RN	(1)		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=86176787; PubMed=2421257;		
RX	Stahl H.-D., Bianco A.E., Crewtcher P.E., Burkot T., Coppel R.L.,		
RA	Brown G.V., Anders R.F., Kemp D.J.;		
RT	"An asparagine-rich protein from blood stages of Plasmodium		
RL	falciparum shares determinants with sporozoites.";		
CC	Nucleic Acids Res. 14:3089-3102(1986).		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
EMBL; M24328; AAA29491.1; --			
-R	PIR: A23770; A23770.		
KM	Malaria.		
FT	NON_TER		
FT	537 537		
SO	SEQUENCE 537 AA; 63000 MW; A6D76B15318CC239 CRC64;		

Query Match 6.6%; Score 203; DB 1; Length 537;			
Best Local Similarity 17.8%; Pred. No. 0.035;			
Matches 93; Conservative 106; Mismatches 212; Indels 112; Gaps 19;			
OY	18	NNNOSINTDDNQIHKTEETNNYDGIERSDETESTTVADENEATFLQKTPQDWTLTE	77
Db	4	NNNNNNNDGN--INYONTNEFKDNKKMNKNKNOYNNYKRDENANNSTTHSRNSAYE	61
OY	78	EEVKSSSVESNSSSID--TAQOPSHTTI-----NREESVOY--SDNVESHVSDEAN	126
Db	62	EHLR--NSSIDNNNSNINNTYNQOTRPSPEMENENENENKNTYHGGMNNNIHFKNKXDNN	120
OY	127	SKIKESNTSGEGEENTIEDPKYKKEBSTISOQSGITINDIKSIODELLNL-PINEYEAK	185
Db	121	SSMK--NDNNKTTDTSYNNKGITINN-----NNMDYLENNINNINEYKGS	163

CC CC

OY 186 ARPSTTGAQPSIRKVTYNALAEAGSGVNHILIKYTDSITIEGDDSGEADKADELII 245
DB 164 AKNKKTYTN-----YMKNKKLKFTQNNDMDANINEDNNNNNNNNNNNGVSNTQNNMNM 217
OY 246 YDYTFEEVDKYKSGDWTMTVIDDKNTVPSPDLTDSTFIPIKONSGEIIATWGYDNKKQIT 305
DB 218 RNSNISIKRNLI-----NNNNIINNMMKMGSODKQNSNNNFYWNYNQNRKSNM 268
OY 306 YTFDTYVKYEIKAHLLTSYIDKSFKPNNNTKLDVEKTAALSSVN----- 352
DB 269 NMMNNNNNNNNHHNNNNNNNNHNNNNNNNNNNNNNNNNNNNNNNNNNNNNINSLSDSMSPNYHAH 328
OY 353 -KTIVFQRPRENTAMIOSM-----FTNIDTKRHTV-----EQRTYLN----- 391
DB 329 VKMSMNMYNN-NESNTAPNQMNFEQTQNDNKRENNMMANNVGYDYNVAHYNNNPSTD F 387
OY 392 -----PLRSAKETNVNI-----SGNGD-----EGSTIIDSTTIKYK-----VG 427
DB 388 FSRAVGNYYNNY-LNNNNNNNAVVNNNSSGMNMKNENSENKVAYADNDSLNNMKNNNNNIN 447
OY 428 DNQNLPDFSRRIDYSEY-----EDVTNDYDAQGLNNNDVIN 464
DB 448 MNESIINNNTLTNNNEYYNNQNNNEDEDDDWGELEGDXYIDIN 490

RESULT 4
HMWL_MYCGE ID HMWL_MYCGE STANDARD; PRT; 1139 AA.
AC Q49413; Q49365;
DI 01-NOV-1997 (Rel. 35, Created)
DI 01-NOV-1997 (Rel. 35, Last sequence update)
DI 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cytoadherence high molecular weight protein 1 (Cytoadherence accessory
protein 1).
GI Hmwl_OG MG312.
OS Mycoplasma genitalium.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
NCBI_Taxid=2097;
RX [1]
RN SEQUENCE FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE=96026346; Pubmed=7569993;
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischmann R.D., Bult C.J., Kierlavage A.R., Sutton G., Kelley J.M.,
RA Frithman J.L., Weisman J.F., Small K.V., Sandusky M., Fuhrmann J.B.,
RA Nguyen D.-T., Dougherty B.A., Bott K.F., Hu P.-C., Lucier T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT "The minimal gene complement of Mycoplasma genitalium.";
RL Science 270:397-403(1995).
RN [2]
RP SEQUENCE OF 721-847 FROM N.A.
CX STRAIN=ATCC 33530 / G-37;
RX MEDLINE=94075230; Pubmed=8253680;
RA Peterson S.N., Hu P.-C., Bott K.F., Hutchison C.A. III;
RT "A survey of the Mycoplasma genitalium genome by using random
sequencing.";
RL Bacteriol. 175:7918-7930(1993)

- I- FUNCTION: COMPONENT OF THE CYTOSKELETON-LIKE STRUCTURE WHICH STABILIZES THE SHAPE OF THE WALL-LESS MYCOPLASMA. THIS CYTOSKELETON-LIKE NETWORK OF ACCESSORY PROTEINS CONTAINING HMW PROTEINS 1 TO 5 ALLOWS THE PROPER ANCHORING OF CYADESIN PROTEINS IN THE MYCOPLASMAL MEMBRANE AT THE ATTACHMENT ORGANELLE (BY SIMILARITY).

- I- SUBCELLULAR LOCATION: LOCALIZES SPECIFICALLY TO THE ATTACHMENT MEMBRANE (BY SIMILARITY).

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: U39712; AAC71534.1; -
 DR EMBL: U02261; AADI2527.1; -
 DR TIGR: MG312; -
 KM Cytochrome c; Structural protein; Complete proteome.
 SQ SEQUENCE 1139 AA; 130531 MW; 0011D3286C3DB856 CRC64;

Query Match: 6.5%; Score 199.5; DB 1; Length 1139;
 Best Local Similarity 21.0%; Pred. No. 0.12;
 Matches 129; Conservative 98; Mismatches 225; Indels 161; Gaps 27;

QY 1 HHHHHHPSDEKNDVNNNO-SINTDNNQI---IKKEETNNYOG-----IE 44
 DB 578 HHEELKPAVEEQNNYOVGDQVQANLDNNEELQPTAEKVTLEDFPSKQAVVDSYOLPID 637
 45 KRSERD-----ESTVNDENATEFLQKTPQDNTHLTFEEVKES-----SVESNN--- 90
 DB 638 TQQOQQTITSSSEFQTPYVEQFDQVNSEVNDQFKPEITKEPVLESFKNQDVETSNTIN 697
 QY 91 --SSIDTAQPSHTTINREESVQTSNDVEDSHVSDFANSKIKES---NTESGKEENTIE 144
 DB 698 NQKFDIOGDNKITITTKKSPQIPTLLPISFVSNRIEKKPVEPLALDNKESQOQOQIT- 756
 QY 145 QNKKKEDSTSQPSGYTRIDE-KISNO---DELNLPIEYENKARPLSTSAQSI 198
 DB 757 --NSTTESKTLAKTSLVOLQINSINNOISYTSSEVRLDKKDDLLTINTVSEDOQPI 814
 QY 199 KRVYTNQLAAEQSNVNLKIKVTQDSTIEGYDSEGVIKAHDAENLYDVFEVDQYKS 258
 DB 815 EYFVNAKEVEEHSITQNKQSVEDKSELDNFKKSDL-----YKISSELR 860
 QY 239 GD-TWTVDDIKATVPSDLTDSFTPIKIDNCGEIIATGYDNKNKQITVTFDYDYKXEN 317
 DB 861 GELNPTINFD-----AIFOMNDYQKSVQSFHLDNPFVNTYKN 898
 QY 318 -----IKAHL-KLTSYIDSKR---VPNNNFK-----LDVEYKRA 347
 DB 899 QISERYLIKKELQSELSLIDQENELNVOFNNAKLLTTLQKEMIRSLASPAIAYKPS 958
 QY 348 -----LSSVNTKIIVEYQRPNNRTANLQSMETNIDTKNHTVEQTI-----Y 389
 DB 959 NSEYDQKSGELMRHYQRAIT-----ENKKTESIQGSLQKQIAYNSCCEIIMNNINKL 1013
 QY 390 TNPRLYSAKETVNVNISGND-----EGSTIIDSTIIKYYKVDNQLDPSNRIYDY 441
 DB 1014 DNTLRFRAKKEKDPDLLSNEDSYTDNGLVENPQMLMD--LIDFSNTEFDN-----ISNEQLDD 1067
 QY 442 SEYEDV-TNDQYALQGNNDVNIINGNDSPYIIIVISKYDKNKDYTIQQTIVT----- 495
 DB 1068 FYIENNDRIIDREFEGFNDN---FYVDIAK-VQMSMAFVYNDLDETLPDRITSNESS 1122
 QY 496 -MOTTINEYTGEEF 507
 DB 1123 LDEDLFESSGDF 1135

RESULT 5
 YMOB_CAEEL STANDARD: PRT; 918 AA.
 AC P34487;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Hypothetical 96.6 kDa protein F59B2.12 in chromosome III.
 GN F59B2.12.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Rhabditinae; Caenorhabditis.
 OC NCBI_TaxID=6239;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Almscough R., Anderson K., Baynes C., Berts M.,
 RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fraser A.,
 RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
 RA Johnston L., Jones M., Kershaw J., Kirsten J., Lalister N.,
 RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders E., Staden R.,
 RA Sims M., Smaildon N., Smith A., Smith M., Sonhammer E., Staeden R.,
 RA Sulston J., Thierley-Mieg J., Thomas K., Vaudin M., Vaughan K.,
 RA Waterson R., Watson A., Weinstock L., Wilkinson-Sproat J.,
 RA Wohldman P.,
 RA "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans".
 RL Nature 368:32-38(1994).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: Z11505; CAAT7581.1; -
 DR PIR: S31132; S31132.
 DR WormPep: F5982.12; CE01024.
 KM Hypothetical protein.
 SQ SEQUENCE 918 AA; 96560 MW; E464FD86B14945DE CRC64;

Query Match: 6.4%; Score 197; DB 1; Length 918;
 Best Local Similarity 20.5%; Pred. No. 0.12;
 Matches 142; Conservative 91; Mismatches 282; Indels 176; Gaps 27;

QY 2 HHHHHPS-----DEKNVNNNOISINTDNNQIIRK----- 33
 DB 74 HKGHQSSGSSNTSLVYVAGDGNITENSCKKQGYKESVYDENENWTIKSADGSYI 133
 QY 34 -----KEETNYDGIKRSERDTE-----STTVNDENEA-----TFLOKT 68
 DB 134 ETGKSHNKSDDASSY-GLEKSKTYADKNGTMTLSNTRKKNQNSALADGNEFVNOQ 192
 QY 69 PODNTHL-----TEEVKESSSVESNSSIDTAQPSHTTINREESVQTSN-VEQS 119
 DB 193 NADGTFLLRNNTGKNTDEHLSHNVLDENNAQMSIG-ADGTSNITNKGIVGSHNAASDA 251
 QY 120 HVSDFANSKIKESNRESGKEENTIEOPNKVKEDSYTSQPSGYTINDEKISNDELNLPI 179
 DB 252 H-----SNFESLDAQGNKKSQNTSKAASASG-----SNADFE-SNLESKLNADG 295
 QY 180 MEYENKARPLSTTS-----AOPSIKRYTVN-----QLAAEQSNVNLKIKYTDQSTIEG 228
 DB 296 TSMNSNTGNFNTSYDKATAEEVMKKNNVADGTSMEKSHAGSNSSKINSASQGSPLS 355
 QY 229 YDSEGVIAKHAENLIIDYTFEVDKYSQGTMYVDIDK----- 268
 DB 356 MWGPNG-IKSHSTSNKTDN-YALDEANQASAGISIEQIKNGQRSINNESSIGSKAESR 412
 QY 269 NTVPSDLDSFTPIKIKDNGEIIATGYDNKNKQITVTFDYKYENIKAHLLKLTYSI 328
 DB 413 NNTADTLDSDVANGTSSSHSKASGSLSDENHNKTHALQASVDEHGKMKHNSIDGSR 472
 QY 339 DSKSVNNNTKLDVETKTA--LSSVN-----KTIYEVQRPNNRTANLQSMETNID 378
 DB 473 NKTGEFSEKSEKSIKNAQDGMQVNVKNDTRNRYTEAEKSALEKNEHKNDGDFKD-- 530
 QY 379 TKNHTVEQTIINPLRYSAKKTENVNISGNDGSGSTIIDSTIIKYYKVDNQLDPSNRI 438
 DB 531 -----BSKGSNSRYN-RTDGSNLAAGSVSVGKGVSSNETIASN-A 571

QY 439 YDYSEYEDVNDYAQGLNNNDVNFNIDSPYIKYISKYDPKDDYTTIQFTVMTQ 498
 DB 572 FNTSDAES-NQFDHLKQXTANGTEITHAK-DS-----KOVAASANKSSIDTSMASVADAKG 625
 QY 499 TINETGFERASVYNTTAFTSSGCGGDDLPPEKTYIKGVWEDVDKQIGNNDNEK 558
 DB 626 NKRYKTSQAADSHD--A1SASS-----DVDAKIVKHA-DNSE 660
 QY 559 PLSNVLTLYTPDGTSGKSVRTDEDKGYQFDG 589
 DB 661 S1SDSSNOTASSEHNDSSKQSEHEKRONADG 691
 RESULT 6
 ID BAF1_YEAST STANDARD; PRT: 731 AA.
 AC P14164;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 01-FEB-1995 (Rel. 31, Last annotation update)
 DT Transcription factor BAF1 (ARS binding factor 1) (Protein ABF1)
 (Bidirectionally acting factor) (SFB-B) (DNA replication enhancer-
 binding protein OBF1)
 GN BAF1 OR ABF1 OR OBF1 OR YKLI12M OR YKL505.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OS Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OS Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OC NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A. Pubmed-2511628;
 RA Diffley J.F.X., Stillman B.;
 RT "Similarity between the transcriptional silencer binding proteins
 ABF1 and RAP1.";
 RL Science 246:1034-1038(1989).
 RN [2]
 RP SEQUENCE FROM N.A. Pubmed-2620828;
 RA Rhode P.R., Sweder K.S., Oegema K.F., Campbell J.L.;
 RT "The gene encoding ARS-binding factor I is essential for the
 RT viability of yeast.";
 RL Genes Dev. 3:1926-1939(1989).
 RN [3]
 RP SEQUENCE FROM N.A. AND SEQUENCE OF 128-154 AND 535-555.
 RA MEDLINE-90076149; Pubmed-2686983;
 RA Halter H., Kaveley B., Vandeckerkhove J., Kiefer F., Gallwitz D.;
 RT "Sequence, expression and mutational analysis of BAF1, a
 RT transcriptional activator and ARS1-binding protein of the yeast
 RT Saccharomyces cerevisiae.";
 RL EMBO J. 8:4265-4272(1989).
 RN [4]
 RP SEQUENCE FROM N.A. AND PHOSPHORYLATION.
 RA MEDLINE-91239512; Pubmed-2034654;
 RA Francesconi S.C., Eisenberg S.;
 RT "The multifunctional protein OBF1 is phosphorylated at serine and
 RT threonine residues in Saccharomyces cerevisiae.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:4089-4093(1991).
 RN [5]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-92221689; Pubmed-1561835;
 RA Jacquier A., Legrain P., Dujon B.;
 RT "Sequence of a 10.7 kb segment of yeast chromosome XI identifies the
 RT APN1 and the BAF1 loci and reveals one tRNA gene and several new open
 RT reading frames including homologs to RAD2 and kinases.";
 RL Yeast 8:121-132(1992).
 RN [6]
 RP SEQUENCE OF 585-731 FROM N.A.
 RA MEDLINE-94344141; Pubmed-8065362;
 RA Butler A.R., White J.H., Fojlawiyo Y., Edlin A., Gardiner D.,
 RA Stark M.J.R.;
 RT "Two Saccharomyces cerevisiae genes which control sensitivity to G1
 RT arrest induced by Kluyveromyces fragilis toxin.";

RL Mol. Cell. Biol. 14:6306-6316(1994).
 CC -1- FUNCTION: TRANS-ACTING FACTOR IN THE REGULATION OF TRANSCRIPTION
 CC AND IN DNA REPLICATION. INVOLVED IN THE TRANSCRIPTION ACTIVATION
 CC OF A SUBSET OF RIBOSOMAL PROTEIN GENES. BINDS THE ARS-ELEMENTS
 CC FOUND IN MANY PROMOTERS. BINDS TO THE SEQUENCE 5'-TCN(7)ACG-3'.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- PM: EXTENSIVELY PHOSPHORYLATED ON SER AND THR RESIDUES.
 CC -1- SIMILARITY: STRONG TO KLUYVEROMYCES ABF1, AND LOCAL TO YEAST RAP1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X16385; CA344421.1; -;
 DR EMBL: X51654; CA35966.1; -;
 DR EMBL: M29067; AA66311.1; -;
 DR EMBL: M63578; AA34823.1; -;
 DR EMBL: S93804; AA22002.1; -;
 DR EMBL: 228111; CA481951.1; -;
 DR EMBL: X77511; CA454647.1; -;
 DR PIR: S29870; S29870.
 DR TRANSFAC: T00056; -;
 DR SGD: S0001595; ABF1.
 DR Transcription regulation; Activator; DNA-binding; Nuclear protein;
 KW Phosphorylation; Zinc-finger; Metal-binding; Zinc; DNA replication;
 KW Trans-acting factor.
 KM ZN_FING
 FT MOD_RES 624 624 PHOSPHORYLATION (BY PKC) (POTENTIAL).
 FT MUTAGEN 57 57 H->Q: LOSS OF DNA-BINDING.
 FT CONFLICT 71 71 C->S: LOSS OF DNA-BINDING.
 FT CONFLICT 125 125 N->K (IN REF. 3).
 FT CONFLICT 128 128 A->V (IN REF. 3 AND 5).
 FT CONFLICT 148 148 I->T (IN REF. 3 AND 5).
 FT CONFLICT 279 280 TN->NT (IN REF. 3 AND 5).
 FT CONFLICT 690 690 N->T (IN REF. 3, 5 AND 6).
 SO SEQUENCE 731 AA; 81749 MW; BF200372F3A9EA3F CRC64;
 Query Match 6.3%; Score 195; DB 1; Length 721;
 Best Local Similarity 20.7%; Pred. No. 0.11; Indels 192; Gaps 33;
 Matches 138; Conservative 95; Mismatches 242;
 QY 1 HHHHHSSDEDEKNDYINNOSIN-----TD-----NNQIKKEETNNYDIEKR 46
 DB 105 HHHHNNNNNNEDND--NNNGSNKASNDKLDFFVDDLEHLANIHPDGTN--DKVESR 160
 QY 47 SEDRTSTTNVDENETFLQKTPQDTHLTLEEVKSSSVESNSSIDTAQPSHTT----- 103
 DB 161 SNE--VNGNNDODADANNIFK--QGGVTIKNDIEDSINKASIDRLDDESGFTGNDSG 216
 QY 104 ---INREESYOT-----SDNVEDSHVS--DFANSKIKESVTESGRENTIEQPNKKEDS 153
 DB 217 NRRHNEEDDVHTOMTKNTSYDYNVEDINVAIAVAANDSQSNKKDG-----KDD 268
 QY 154 TTSQPSGYTNIDEKISNOELLNLPINEYENKRAPLSTSAQPSIRYVNOIAAOGSN 213
 DB 269 ATNNNGGQDN-----TNNNNHNNNSNIN--NNNGSHGISHSPPSIRDSKNL----- 315
 QY 214 VNHLLVTDQSTEGYDSEGVIAKHAENLIYDFEVDDKYSQDYMVTVIDKNTVPS 273
 DB 316 -----DVFNASATDIDPGPFVVTIKIEPY--HSHPLEDNLSTLKF----- 352
 QY 274 DLTDSFTIKIKNSGEI--IATGYDNKNKQIT--YTFDYDYDK-----YE 316
 DB 353 -LT--KIPRILONDLKFDQILLESSTNNSHVSKFVSHVYESGLDLTKMORYLTME 408
 QY 317 NIKAKL-----KLTSY-----IDRSKVPNNNTKLVEYKTKALSSVN--K 353
 DB 409 DFEKRLLSQIARITTYIKKAFVLKKKKMGREYNLDQSSSSNNNNNNNDGE-----LSGTNLR 464

[illegible]

OC Candida albicans (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
 NX NCBI_TaxID=5476;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 10261;
 RX MEDLINE=96133936; PubMed=8552638;
 RA Gale C., Finkel D., Tao N., Meinke M., McClellan M., Olson J.,
 RA Kendrick K., Hostetter M.;
 RT "Cloning and expression of a gene encoding an integrin-like protein
 in *Candida albicans*.";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:357-361(1996).
 CC -1- FUNCTION: COULD PLAY A ROLE IN ADHESION AND IN STE12-INDEPENDENT
 CC MORPHOGENESIS.
 CC -1- SUBCELLULAR LOCATION: CELL SURFACE OF THE BLASTOSPORES.
 CC -1- SIMILARITY: SOME, WITH INTEGRIN ALPHA CHAINS.
 CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U35070; AAA96019.1;
 DR InterPro: IPR001849; PH.
 DR Pfam: PF00169; PH.1.
 DR SMART: SM00233; PH.1.
 DR PROSITE: PS50003; PH_DOMAIN; 1.
 DR Cell adhesion; Glycoprotein.
 KW DOMAIN 1527 1636
 FT SITE 1149 1151 PH.
 FT CARBOHYD 50 50 CELL ATTACHMENT SITE (POTENTIAL).
 FT CARBOHYD 154 154 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 216 216 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 223 223 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 265 265 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 268 268 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 285 285 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 289 289 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 306 306 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 314 314 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 446 446 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 583 583 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 638 638 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 670 670 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 691 691 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 817 817 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 838 838 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 841 841 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 1018 1018 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 1082 1082 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 1100 1100 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 1103 1103 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 1113 1113 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 1200 1200 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 1571 1571 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 1593 1593 N-LINKED (GLCNAC...) (POTENTIAL).
 FT DOMAIN 95 106 POLY-GLN.
 FT DOMAIN 283 286 POLY-ASN.
 FT DOMAIN 1283 1292 POLY-HIS.
 FT DOMAIN 1651 1660 POLY-GLN.
 SO SEQUENCE 1664 AA; 187859 MW; BA2EE0DC8196790 CRC64;

Query Match 6.1%; Score 189.5; DB 1; Length 1664;
 Best Local Similarity 20.4%; Pred. No. 0.49;
 Matches 131; Conservative 105; Mismatches 236; Indels 171; Gaps 31;

15 DIVINNOQS---INTDNNQIIRK-----EFNNVNGIKRSRSDRESTT 55

Db 138 DKVNNHAPTYINTSPKNSIMKKATPKASPKVATVTYNPEIHNP--DNKVEEDQSOQ 195
 QY 56 NVDENAEATFLOKTPQDNHNLTEEEVKSSVESSNSIDTAQOQPSHTT-----INNE 108
 Db 196 KEDSEYEPPLIQ-----HOKMDPSQFNYSDEEDTNAASVPT-PLHTITKPPFAQLNN 247
 QY 109 SVQISDNVEDSHVD-----FANSKIKE-----STESGKEENTIEQPKKEDSTT 155
 Db 248 EV---NSEPEALTDMLKRENFNSLSEDEKYNILSTPTNNNSKNSKVDMSHO-----N 299
 QY 156 SQPSGYTINDEKISQDDELINPIVEYENKARPL-STISAPSTIRRVYNQDLAAGSNV 214
 Db 300 LODASKRNTNENIHNLSFALAPKNDIEN--PLNSLTNADISLSKSSGSSQSLQSLAND 356
 QY 215 NHLIV---TDOSTIEGYDSDSEGVYKADAEVLYDYFVEVDKVSQDFTMTVIDKNTV 271
 Db 357 NRVLESVSGSPKKNVPGSLNDGKGFSD-----EVYESLL 392
 QY 272 PSDLT-DSFTTPKIND-----NSGEIATGYDNKNKQITTYTVDYKYNIKAH---- 321
 Db 393 PDLRSRDLKLETKEDHDAPEHNNEFIDAKSTNTNG-OLVSSDDHLSFDRSYNHTQOS 451
 QY 322 -LKLTSTYIDKSKVPMN-----NPKLDVYKTAALSSVYNK 353
 Db 452 ILLNLSASQSOISLNALEKORQOEOQOAAEPEEETSFSDNKTIVQEPKSNLEFKV 511
 QY 354 TITVEQRPNERNTANLSMFTNIDTKNH-TVEQTYINPLRYSAKETNVNISG----- 406
 Db 512 TIKKEPVASATEIKAKRRESSRLIKNDELAEPADIHKKENANSHVEDTALLKA 571
 QY 407 -NGDEGSTIIDSTTIKY-YKVGDNQNLPSNRITYDSEYEDVYNDYQAQGN--NDV 461
 Db 572 LNDDESDPTTONSTKMSIRFHIDSDMKLEDSND--GDRED--NDDISREFKSDILND 625
 QY 462 N-----INGNDISPIYIKYISKYDKPKDYTTIQGYVMQOTINETGFR-----T 509
 Db 626 SOTSIIIGDKYGNSSSEITTKTLA--PPRSNNDKNSKSLSDPANNSLQQLVPHPT 682
 QY 510 ASYDNTIAFSTSSGOGGDLPPER-----TYKIGDY-VVEDVDR 547
 Db 683 KEDDSILANSSNIA-----PPEELTLPEVEANDVSSFFNDYRK 719
 RESULT 9
 ID EBA1_PLAFC STANDARD; PRT; 1435 AA.
 AC P19214;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE Erythrocyte-binding antigen EBA-175.
 OS Plasmodium falciparum (isolate Camp / Malaysia).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
 NX NCBI_TaxID=5835;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=90377299; PubMed=2204835;
 RX Sim B.R.L.;
 RA "Sequence conservation of a functional domain of erythrocyte binding
 RT antigen 175 in *Plasmodium falciparum*."
 RL Mol. Biochem. Parasitol. 41:293-296(1990).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X52524; CA36756.1;
 DR PIR: S11561; S11561.

DR EMBL; U66740; AAB18388.1; JOINED.
DR EMBL; U66741; AAB18388.1; JOINED.
DR EMBL; U66742; AAB18388.1; JOINED.
DR EMBL; U66743; AAB18388.1; JOINED.
DR EMBL; U66745; AAB18388.1; JOINED.
DR EMBL; U66746; AAB18388.1; JOINED.
DR EMBL; U66747; AAB18388.1; JOINED.
DR EMBL; U66748; AAB18388.1; JOINED.
DR EMBL; U66749; AAB18388.1; JOINED.
DR EMBL; U66750; AAB18388.1; JOINED.
DR EMBL; U66751; AAB18388.1; JOINED.
DR EMBL; U66752; AAB18388.1; JOINED.
DR EMBL; U66753; AAB18388.1; JOINED.
DR EMBL; U66754; AAB18388.1; JOINED.
DR EMBL; U66755; AAB18388.1; JOINED.
DR EMBL; U66756; AAB18388.1; JOINED.
DR EMBL; U66757; AAB18388.1; JOINED.
DR EMBL; U66758; AAB18388.1; JOINED.
DR EMBL; U66759; AAB18388.1; JOINED.
DR EMBL; U66760; AAB18388.1; JOINED.
DR EMBL; U66761; AAB18388.1; JOINED.
DR EMBL; U66762; AAB18388.1; JOINED.
DR EMBL; U66763; AAB18388.1; JOINED.
DR EMBL; U66764; AAB18388.1; JOINED.
DR EMBL; U66765; AAB18388.1; JOINED.
DR EMBL; U66766; AAB18388.1; JOINED.
DR EMBL; U66768; AAB18388.1; JOINED.
DR EMBL; U66769; AAB18388.1; JOINED.
DR EMBL; U66770; AAB18388.1; JOINED.
DR EMBL; U66771; AAB18388.1; JOINED.
DR EMBL; U66772; AAB18388.1; JOINED.
DR EMBL; U66773; AAB18388.1; JOINED.
DR EMBL; U66774; AAB18388.1; JOINED.
DR EMBL; U66776; AAB18388.1; JOINED.
DR EMBL; U66777; AAB18388.1; JOINED.
DR EMBL; U66778; AAB18388.1; JOINED.
DR EMBL; U66779; AAB18388.1; JOINED.
DR EMBL; U66780; AAB18388.1; JOINED.
DR EMBL; U66781; AAB18388.1; JOINED.
DR EMBL; U66782; AAB18388.1; JOINED.
DR EMBL; U66783; AAB18388.1; JOINED.
DR EMBL; U66784; AAB18388.1; JOINED.
DR EMBL; U66785; AAB18388.1; JOINED.
DR EMBL; U66786; AAB18388.1; JOINED.
DR EMBL; U66787; AAB18388.1; JOINED.
DR EMBL; U66788; AAB18388.1; JOINED.
DR EMBL; U66789; AAB18388.1; JOINED.
DR EMBL; U66790; AAB18388.1; JOINED.
DR EMBL; U66791; AAB18388.1; JOINED.
DR EMBL; U66792; AAB18388.1; JOINED.
DR EMBL; U66793; AAB18388.1; JOINED.
DR EMBL; U66794; AAB18388.1; JOINED.
DR EMBL; U66795; AAB18388.1; JOINED.
DR EMBL; M59892; AAB18388.1; JOINED.
DR PIR; A35899; M59892.
DR HSP; P02468; 1KLO.
DR MIM; 156225;
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001886; LamNT.
DR InterPro; IPR000034; Laminin_B.
DR InterPro; IPR002049; Laminin_EGF.
DR InterPro; IPR001791; Laminin_G.
DR Pfam; PF00052; Laminin_B; 2.
DR Pfam; PF00053; Laminin_EGF; 15.
DR Pfam; PF00054; Laminin_G; 5.
DR Pfam; PF00055; Laminin_Nterm; 1.
DR PRINTS; PR00011; EGF-LAMININ.
DR ProDom; PD002082; LamNT; 1.
DR ProDom; PD003031; Laminin_B; 2.
DR SMART; SM00180; EGF_Lam; 14.
DR SMART; SM00001; EGF_like; 3.
DR SMART; SM00281; LamB; 2.

DR SMART; SM00282; LamG; 5.
DR SMART; SM00136; LamNT; 1.
DR PROSITE; PS00022; EGF_1; 11.
DR PROSITE; PS01186; EGF_2; 3.
DR PROSITE; PS01248; LAMININ_TYPE_EGF; 14.
DR PROSITE; PS50025; LAM_G_DOMAIN; 5.
KW Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;
Laminin EGF-like domain; Cell adhesion; Repeat; Signal; Polymorphism.
FT SIGNAL 1 22
FT CHAIN 23 3110 LAMININ ALPHA-2 CHAIN
FT DOMAIN 23 286 LAMININ N-TERMINAL (DOMAIN VI).
FT 287 343 LAMININ EGF-LIKE 1.
FT DOMAIN 344 413 LAMININ EGF-LIKE 2.
FT DOMAIN 414 468 LAMININ EGF-LIKE 3.
FT DOMAIN 469 517 LAMININ EGF-LIKE 4.
FT DOMAIN 518 527 LAMININ EGF-LIKE 5 (N-TERMINAL).
FT DOMAIN 528 723 LAMININ DOMAIN IV 1 (DOMAIN IV B).
FT DOMAIN 724 756 LAMININ EGF-LIKE 5 (C-TERMINAL).
FT DOMAIN 757 806 LAMININ EGF-LIKE 6.
FT DOMAIN 807 864 LAMININ EGF-LIKE 7.
FT DOMAIN 865 917 LAMININ EGF-LIKE 8.
FT DOMAIN 918 966 LAMININ EGF-LIKE 9.
FT DOMAIN 967 1013 LAMININ EGF-LIKE 10.
FT DOMAIN 1014 1059 LAMININ EGF-LIKE 11.
FT DOMAIN 1060 1105 LAMININ EGF-LIKE 12.
FT DOMAIN 1106 1165 LAMININ EGF-LIKE 13.
FT DOMAIN 1166 1175 LAMININ EGF-LIKE 14 (N-TERMINAL).
FT DOMAIN 1176 1379 LAMININ DOMAIN IV 2 (DOMAIN IV A).
FT DOMAIN 1380 1419 LAMININ EGF-LIKE 14 (C-TERMINAL).
FT DOMAIN 1420 1468 LAMININ EGF-LIKE 15.
FT DOMAIN 1469 1526 LAMININ EGF-LIKE 16.
FT DOMAIN 1527 1573 LAMININ EGF-LIKE 17.
FT DOMAIN 1574 2144 LAMININ II AND I.
FT DOMAIN 2145 2328 LAMININ G-LIKE 1.
FT DOMAIN 2340 2521 LAMININ G-LIKE 2.

Query Match 6.0%; Score 186.5; DB 1; Length 3110;
Best Local Similarity 19.9%; Pred No. 1.3; Mismatches 256; Indels 215; Gaps 32;
Matches 142; Conservative 102;

QY 28 NNOIIRKEET-----NNTDGIKRSEDRTESTNVNENAFLOQTPDQNTLHTEE 78
DB 1693 NEKAIKLTETLTGRDEAFERNLEGLQK-----EIDQIKELRRKILTEQREIAD 1742

QY 79 EYKSSSVSSNSSIFPAQPSHTTIN---REESVQTSNVEDS-HVSDPAISKIESN- 133
DB 1743 ELVAEALIKKVKKLPGESGENSEMEKDLREKLADYKKNVDAMDLREATDKIREANR 1802

QY 134 -----TESGKE--ENTIEOPNKV-----KEDSTS 156
DB 1803 LPAVNOKNMTALEKKKEAVESGRQLENTLKBENDILDEANRLADEINSIIVDEIQT 1862

QY 157 OPSGYTINDEKISN---ODELLNPINEYENKARPLSTSA---OFSIKRVTVNOL 206
DB 1863 LPPMSEELNDKIDDLDSQEIQRKLAKVSGAESHAQNDSSAVLDGLIDEAKNISFNAT 1922

QY 207 AAEQG-SVNNHLIKVTDOSTTSGYDDESEGYKADNENLIYDFEYVDKVGSGDMTVD 265
DB 1923 AAFKAYSNIKDYI-----DEAEKVAK--EAKDLAHEAT-----KLAGPR--G 1961

QY 266 IDKNTPPSDLTDSFTI-----PKIDNSGEIITATGT-----YDNKNKQITVTFDYVD 313
DB 1962 LKEDAKKGLCKLSFRILNEKKLANDEVKEDHNLGKTRIEVADARNGLDLTLMDTGL 2021

QY 314 KYENI--KAHLKLTXYIDSKVYNNNTK-----LDVEYKTAISSVKKTT 356
DB 2022 KLSAIFNDPAKKAQAVKDKARQANDTAKDYLQITELHQNLGLKKRYNKLADSVAKTA 2081

QY 357 VEYQRPENR--TANLOSMFTNID-----TKNNTVEQTITINPLKRS 396
DB 2082 V-VKDESKNKTITADADATVNLDEQADRDLIDKLPIKELEDNKKNISSEIKELINOARQ 2140

QY 397 AKETVNVISGND-----EGS-----TIIDSTIIKV-----YK 425

```

DB 2141 ANSIRKVSAGGDCITRTYKPEIKKSYNNIYVAVKTAADVADNLLFYGSAKFIQFLAEMR 2200
QY 426 VGDNQNLPPSNRIYDSEEDVTND-----YAQLGNNDVNIINFGIDSPYIIKIVTS 478
DB 2201 KGVKSTLMDVGVSGVCHVEPDLLIDDSYVRIYASRTGRNGTISR--ALDGPKASIVPS 2258
QY 479 -KXDPKKDDYTTIQVTVMTQITNETGERTASYNNTAFSSSGGGDLPPEKTYKI 537
DB 2259 THHSTSPGGTITLDVANNMLFVGLTGKLRKADAVRTFTSCMETYFDNKP-----I 2313
QY 538 GBYVMEVDKDGIONTNDNEKPLSNVLTLYPD--GTSKSVRT--DEGKYOPDG 589
DB 2314 G--LWNRREKES-----DCKGCTVSPRYVDESEGTIOFDG 2345

RESULT 11
YBET_YEAST STANDARD; PRT: 1381 AA.
P34216;
01-FEB-1994 (Rel. 28, Created)
01-OCT-1994 (Rel. 30, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Hypothetical 150.8 kDa protein in SEC17-QCRL intergenic region.
YBL047C OR YBL0520 OR YBL0501.
Saccharomyces cerevisiae (Baker's yeast).
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomycos.
NCBI_TaxID=4932;
[1]
SEQUENCE OF 1-961 FROM N.A.
STRAIN=5288C;
MEDLINE=95116707; PubMed=7871888;
de Wergifosse P., Jacques B., Joniaux J.-L., Purnelle B., Skala J.,
Coffeau A.,
"the sequence of a 22.4 kb DNA fragment from the left arm of yeast
chromosome II reveals homologues to bacterial proline synthetase and
murine alpha-daptin, as well as a new perase and a DNA-binding
protein.";
Yeast 10:1489-1496(1994).
[2]
SEQUENCE OF 579-1381 FROM N.A.
STRAIN=5288C;
MEDLINE=94205266; PubMed=8154187;
Scherens B., el Bakhouy M., Vierendeels F., Dubois E., Messenguy F.;
"sequencing and functional analysis of a 32,560 bp segment on the
left arm of yeast chromosome II. Identification of 26 open reading
frames, including the KIP1 and SEC17 genes.";
Yeast 9:1355-1371(1993).
-1- SIMILARITY: BELONGS TO THE VDP/USO1/YBL047C FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----
CC EMBL: 235808; CAAB4867.1;
CC EMBL: X78214; CAAS5048.1;
CC EMBL: 223261; CAAB0797.1;
CC PIR: S45781; S45781.
CC SGD: S0000143; YBL047C.
CC InterPro: IPR000261; EFS15_repeat.
CC InterPro: IPR000449; UBA.
CC Pfam: PF00036; ehand; 2.
CC SMART: SM00054; EPh; 2.
CC SMART: SM00027; EH; 3.
CC SMART: SM00165; UBA; 1.
CC Hypothetical protein.
CC SEQUENCE 1381 AA; 150783 MW; 626FD26IDCBA7D99 CRC64;

```

```

Query Match 6.0%; Score 186; DB 1; Length 1381;
Best Local Similarity 20.0%; Pred. No. 0.55;
Matches 143; Conservative 115; Mismatches 262; Indels 194; Gaps 34;

QY 7 PSSDEKNDYINN-NQSIMTDDNN--QIIKKE-----TNNYDIKRSSEDRSESTN 56
DB 437 PSPTKQTVVQNNNTNNSFYDNNNGATLQQQPPQPPPLTHSSSGIKFT-----PTSN 491
QY 57 VDENEATFLPKTPODNTHLTEEE-----VKSSSVSSSSSIDTAQOPSHTTINR 106
DB 492 FQD---SIKKEPEBEQRLRESSDTSAPPPPKHASSPVKRTATTLTPQVNFVFSM 548
QY 107 -----EESVOTSD-NVEDSH 120
DB 549 PAGAATSAATGAAGAAGAAALGASAFSSRNAKKNODLFADGEASQLSNATTEMAN 608
QY 121 VSPFANSKIKESNTESGKEENTIEOPNKYKEDSTGSPSGYNI-----DEKISNDEL-- 174
DB 609 LSNQVNSLSKQASITDCKSRATQELKRTTEMKNSIQIK-LNNLRSTHDQNVACTOLEA 667
QY 175 LNLPIIN-EYENKARPLSTTSAPPSIKRVYVNOALABOGSNVHLIKVTDQSITEGYDSE 233
DB 668 QVLQVKNENETLAAQLAVSEA-----NYHAAE--SKNELTDLQESQTKNAELKE 716
QY 234 GYIKAHDAENLIDYVFEVDKRYKSGDTMTVDIDKNTVPSDLTDSFTIKRINSGEIIA 293
DB 717 QITNLMSMTASLOSQLEKQOQVKQERSM-VDVNSKQLELQVTVANLQKEIDGLEKIS 775
QY 294 TGYDNKRNKQITVTFDYVDKYENIKAHKLTSYIDSKVPPNNNTKLDEYKATLSVYNK 353
DB 776 --VYLRKQKEL--NDQKYVEQHAQQL-AKYQDLS---KNDTDL-IDREKQLEERNR 824
QY 354 TTFVEYORPENENTANLQSMFTNIDYKNHTEQT-----IYNPIAR-YSAKETNVN 403
DB 825 QIE-EOENLYHQVSKIQEWFDDLQSKKASFEKADQELKERNIEYANNVRELSEERQMLA 863
QY 404 ISGNDEGSTII-----DSTIIK-VYVGVGNQNLPPSNRIYDVS 442
DB 864 MGQLPEDARDIILAKSASNTDTTKEATSRGNHEDTVSKFETTVENSNL-NNRYKXDE 942
QY 443 E-----YEDVTNDYVQQLGNND--VNINFG-----NIDSPYIIKIVIS-KYDPNKDY- 487
DB 943 EKTERESDVPFDVDTYLSQSDESNANTRNGSQSGNEANPMLTETLSDFRGDLNEXG 1002
QY 488 -----TTIQQVTVMTQITNETGERTASYNNTAFSSSGGGDLPPE 531
DB 1003 IPRQSLTSSVANNAPQSVRDVLELPETLEERTINNTANRQNTGNLSHIPGE----- 1055
QY 532 ETKYKIGDYVMEVDKDGIONTNDNEKPLSNVLTLYPDG-TSKSVRTDEDEK 584
DB 1056 -----WEATPATASTDVLNNE-----TTEVIEDGSTTKRANSNDEGE 1092

RESULT 12
YNP4_YEAST STANDARD; PRT: 1165 AA.
P53950;
01-OCT-1996 (Rel. 34, Created)
01-OCT-1996 (Rel. 34, Last sequence update)
01-OCT-1996 (Rel. 34, Last annotation update)
Hypothetical 128.1 kDa protein in OMP2-MG65 intergenic region.
YND054W OR N2467 OR YNL2467W.
Saccharomyces cerevisiae (Baker's yeast).
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomycos.
NCBI_TaxID=4932;
[1]
SEQUENCE FROM N.A.
STRAIN=5288C / FY1676;
MEDLINE=96021608; PubMed=8533472;
Berger P., Dolignon F., Crouzet M.;

```



```

OY 190 STTSAPSI-----KRVYNOLAEGSNVNHILKTYDOSI---TEGY----- 229
DB 1120 HHADORSVNDITIDEKLSERIDNELAT---VDVNSLANKEQVLDLFEYVDVNV 1176
OY 230 ---DDESEGIKAH-----DAENLIYDVTFE-----VDDKVKSGDTMTVDI 266
DB 1177 GMQDDDSOGYQKEDLPFDGNNILIEIKIQLQTSLSNQEICERVDVDEDI--SGEAKNESV 1235
OY 267 DKNVTP-----SDLTDS---FTIPKIDNGSEIILATGYDNKNKOITYTF 308
DB 1236 EKNQVVDVLPYPAKVYGDQISPLQDEKINLETMETDKNDQGL-----CLEKENE----- 1285
OY 309 TDYVAKYENIK-----AHKLKLSYIDKSKVP-----NNNKKL-----DVEKTKLSS-- 350
DB 1286 TEIIEVTSQFQPADLSHDAGREL--VDQSANLQFCENPKPLIAHHIELEYADSDL 1343
OY 351 -----VAKTIVEYORPENTANLQSMFTNIDTKNHTVEQTIYINPLKYSAKETNVNI 404
DB 1344 ESTEBOYQETERIPF-KPEDSKMEN-----ENSESEBSVDSQELSLN---SHKSEPEI 1393
OY 405 SGNSEGSSTIIDSSTIIKYKVGNDQNLPSNRIYDY-SEYEDVTN--DDYAQLGNNDV 461
DB 1394 S-----KDYOL--EQTLPDVPLPNELEFEDLCEQPDVHEHQNNDS 1435
OY 463 NINFGNIDSPYIKYISKYDPKMDYTIQCVTMQTTINET--GEFRASYNITAFST 520
DB 1436 GA-----STFTYV-----DEKEREERESVSDEESNEEFQDVLSVDKTSQVEVTT 1483
OY 521 SSGQGGDLPEPKYTKIGDYWEDVDKQIQNTND--NEKPLSNV----- 564
DB 1484 LSGLMQ-----EPST-LQD--NESESDSMENAEILNENPSNDLYDFMWSQMTETKIITA 1534
OY 565 -----VTITVPDGTSK---SVRTDEGKYQFD 588
DB 1535 EQVTEQTEVTLQFDDAPNKLFTENLNAKRETYDYE 1569

RESULT 14
YGA4_YEAST STANDARD: PRT: 817 AA.
AC P46949;
DT 01-NOV-1995 (Rel. 32, Last Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical 90.8 kDa protein in RRP41-SNG1 intergenic region.
YGR196C OR G7589.
Saccharomyces cerevisiae (Baker's Yeast).
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=96076633; PubMed=7502584;
RA Guerrero I., Maiz e Silva A., Barreiros T., Arroyo J.,
RA Garcia-Gonzalez M., Garcia-Saez M.I., Rodrigues-Pousada C.,
RA Nombela C.;
RA "The complete sequence of a 9000 bp fragment of the right arm of
RT Saccharomyces cerevisiae chromosome VII contains four previously
RT unknown open reading frames."
RL Yeast 11:1087-1091(1995).

-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X82775; CAA58019.1;
DR EMBL; Z72981; CAA97222.1;

```

```

DR SGD: S0003428; YGR196C.
KW Hypothetical protein.
SO SEQUENCE 817 AA; 90797 MW; E52C5D659D63BEBB CnC64;

Query Match
Best Local Similarity 21.0%; Pred. No. 0.36;
Matches 134; Conservative 88; Mismatches 237; Indels 180; Gaps 31;

OY 8 SDEEKNDVYNN-----NQSTINDNNQI--IKKEETNNYDGIKRSKD 49
DB 120 SEHSKSDYLSASTASLAKSPSEKSKSPHTNRVAVNEDLDNLIEQISREMPF---IRQTSDF 176
OY 50 RTESTTNDE--NEATFLQKTPQDNTHLTEEVYKSSSVSSNSSIDPAQPSHTTINE 107
DB 177 RRDS--DSCDEIQNEAPLPEAPVSSSPVEDEKSHSLGSMQTNADTFE--NTPTRNGN 233
OY 108 ESVQTSNVVEDSHVDPAFNSKIKESN-----TESGKEENTIPQPKVKEEDSTSPSGYT 162
DB 234 EHLSSDGVSEKQKDEF---KYSERGIADILPAKKEEN-LQO-----EDDEVESSGAL 284
OY 163 NIDEKISNDELNLNPINEYENKARPLSTSAQPSI---KRVYNOLAEGSNVNHILK 219
DB 285 EKKESEKETSIRN-----RN-----STSSGQDKYAKPKPVANETKTSIDNGYRMSFMD 333
OY 220 VTDSITEGYDSEGVIAKHAENLIYDTEFVDKVKVSGDTMTVDIKNTVPSDLTD-- 277
DB 334 YQHSSEDEEDNNGENSSSD-----DNRSSVSQKADHINKQQLDTTDD 381
OY 278 --SFTIPKIDNGSEIILATGYDNKNKOITYFTDYVQKYNIAHLKLSYIDKSKYPN 335
DB 382 ALSTY-ESIKYSTNTEEDNEDNESIE-----DKNE-----DNESIED 419
OY 336 NNTKLDEYKATLASSVNTTIVEYQRPNE-----RTANLQSMFTNIDTKN 381
DB 420 ENEDTD-SYKFSNRKGSILTSDEEEKGMSDSEGLKAPKSGYFSKMGN--DKG 477
OY 382 HTVEQTIYINPLKYSAKETNVNISGNGDEGSTIIDSSTIIKYKVGNDQNLPSNRIYDY 441
DB 478 DSALQPNQIDRT---ENTLNSNSGSELENSDGSDEEHINEDKYLEESSVVDSTVDVSW 533
OY 442 SEYEDVTNDYKQLGNNDVNINFGNIDSP--YII-----TVISKYD-- 481
DB 534 KPDSEALSSGFVQ-----DTANKKAPPGIVLDSNKLVDLPASAKPRVYSIYSMA 584
OY 482 -----PNK---DDYTIQQTVMQTTINETYGEFRASYNITAFSTSSQG--QGDLP 530
DB 585 ESTWDAFPSKGEDDLLETIRDTKI-----YNNNTIYVNPGLIGNSNLP 629
OY 531 P-----EKTYYKIGDYWEDVDKQIQNTNDNKPPLSNVL 564
DB 630 PLPMDAQEQOLNAGN-----DNSTTDNDSNNTNANDL 660

RESULT 15
RBP2_PLAVB STANDARD: PRT: 1251 AA.
AC Q00799;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Reticulocyte binding protein 2 (Fragment).
GN RBP2.
OS Plasmodium vivax (strain Belen).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=31273;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=92315338; PubMed=1617731;
RA Galinski M.R., Medina C.C., Ingravallo P., Barnwell J.W.;
RA "A reticulocyte-binding protein complex of Plasmodium vivax
RT merozoites."
RL Cell 69:1213-1226(1992).

```

Search completed: July 30, 2002, 10:41:04
Job time: 548 sec

```
CC -1- FUNCTION: INVOLVED IN RETICULOCYTE ADHESION. SPECIFICALLY BINDS TO
CC HUMAN RETICULOCYTE CELLS.
CC -1- SUBCELLULAR LOCATION: Membrane-bound (Probable).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M88098; AAA29744.1; .
KM Malaria; Receptor; Membrane.
FT NON_TER 1251 1251
SQ SEQUENCE 1251 AA; 143741 MW; 54BA51C7404AC572 CRC64;

Query Match 6.0%; Score 184; DB 1; Length 1251;
Best Local Similarity 20.0%; Pred. No. 0.59;
Matches 140; Conservative 115; Mismatches 236; Indels 210; Gaps 37;

UY 18 NNNOSINTDDNNOITKKEE--TNNYDGIEKRSERDPE-----STTNV--DENEAPELQK 67
DB 302 SSSNKNVINEINENINRSQYIKDIEDAKQASTKVELFKHETTSNIFKESSEILGVET 361
OY 68 TPQDNTHLTEEVKSSSVSSSSSIDTQAQPSHTTINREESVQSDNVEDSHVSD-FAN 126
DB 362 KQOKINKKAEIDIMKE--IERHNSIEIQGVKQFQENLKNLNEPHNYDNMADELNDKSTN 418
OY 127 SKIK-ESNTESGKEENTIQPKVKEDSTTSQPSGTYNIDEKISHODELLNPINEYENK 185
DB 419 AKVLIETNLESVK--HNLESEITNIKQGG-----EKIYSK-----AKDIMOK 457
OY 186 ARPLSTTSQPSIKRYTVNQLAEOGSNNHLIKVTDOS---ITE-----GYDDS---E 233
DB 458 IKATSENTAKETLEKY-----KDDOSNYVNYLQITTEKRLNLTGKRLNGIDSTITNE 512
OY 234 GVIAKHAENLIYDVF-EVDDKVKSGDTMYDIDKNTVP-----SDLTSEFTIPKID 286
DB 513 GALKESKGN---YEIGLEKLEIGKRNRLKYDITKKSINSFVGNFSSLFNFDLNOYDF 569
OY 287 NGEIATGTQYONKNQIYTFIDYDXY-ENIK-----AHLKLTSYIDSKY 333
DB 570 NKN---INDYENKKEIYNEFEGLNKISENLNASENTSDYNSAKTLRLLEQKERVNL 625
OY 334 PN-----NNTKLDVE-----YKTAASSVKNKTITVEYQRPNNRPTANLQSMFTNID 378
DB 626 LKKEEANKYLRVKVESFRPIFNKKESLDKINEMIKKEQLVNEGH-GNKKQVLENIK 684
OY 379 T-----KHHVTEQTIYINPLRYSAKETNVNISGNDEGSTIIDSTTIK 422
DB 685 ELVDENNLSDILKQATGKNEEI-QKITHSTLKNKAK---TILGHVDTSAKYVG---IK 735
OY 423 V-----YKVGDNQNLPSNRIVYSEYEDVTNDYQALGNNNDVNNINFGNIDS 470
DB 736 IPELALTELGLDAKLKTAQELFEKSKNNVLTENNKK-----NTNELDVH-KNIQD 787
OY 471 PY--IIKVISKYD-----PNKDDYTTIO--QVTWQTTINEYTGGEFRT 509
DB 788 AYKVALEIILAHSEIDITKQDSSKLIEMNQIYLKAVYLINQYKKNKISSISKBEAVSVKI 847
OY 510 -----ASYDNTIAF--STSSGQGGDLPEPEKTYKIGDYWEDVDKGI 550
DB 848 GNVSKKHSELSKITCSGSKSYDNIIALEKQELQNLNRSFTQEK----- 891
OY 551 QNTNDNEK-----PLSNVLVTLTYPDGTSKSVRTDED 582
DB 892 -NTNSDSKLEKIKTFESLKNALKTL--EGEVNALKASSD 928
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 30, 2002, 10:31:01 ; Search time 88.07 Seconds

(without alignments)
1164.823 Million cell updates/sec

Title: US-09-147-405-11

Perfect score: 3087
Sequence: 1 HHHHHHPSDEKNDVNNNN.....SKSVRTDEGKTPGQVOD 593

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SPREMBL_19:*
2: sp.archaea:*
3: sp.bacteria:*
4: sp.fungi:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.podent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*
15: sp.virus:*
16: sp.bacteriaph:*
17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3015	97.7	1092	2	070022
2	2812.5	91.1	931	2	09K113
3	1162.5	37.7	1171	2	09KWX6
4	1146.5	37.1	1166	2	086489
5	1139.5	36.9	1141	16	099W46
6	1139.5	36.9	1141	16	093287
7	576	18.1	1315	2	086488
8	559	18.1	1385	16	099W47
9	506	16.4	1733	2	09K114
10	490	15.9	953	16	099W48
11	460.5	14.9	881	2	093M7
12	456	14.8	947	2	086487
13	452	14.6	933	2	053653
14	452	14.6	935	16	0932C5
15	452	14.6	989	16	099V14
16	405.5	13.1	877	16	099R07

17	401.5	13.0	913	2	086476	086476 staphylococ
18	376	12.2	940	2	053682	053682 staphylococ
19	345.5	11.2	961	16	099RD3	099RD3 staphylococ
20	326	10.6	1038	16	099RD2	099RD2 staphylococ
21	273.5	8.9	2402	2	09AER7	09AER7 staphylococ
22	261	8.5	1161	2	09X3M7	09X3M7 streptococ
23	239.5	7.8	2276	2	093TV6	093TV6 staphylococ
24	239	7.7	1039	2	P72534	P72534 streptococ
25	234	7.6	604	5	Q26021	Q26021 plasmodium
26	231.5	7.5	2647	5	Q9U4X0	Q9U4X0 plasmodium
27	225	7.3	1463	2	086919	086919 staphylococ
28	222	7.2	1711	5	077322	077322 plasmodium
29	220	7.1	2843	5	Q96315	Q96315 dictyostell
30	218.5	7.1	1938	12	Q9DH49	Q9DH49 amastaxa moo
31	218	7.1	1928	5	Q9U0H2	Q9U0H2 plasmodium
32	218	7.1	2081	10	Q9LH98	Q9LH98 arabidops1s
33	217.5	7.0	559	5	Q9U3Y8	Q9U3Y8 plasmodium
34	216.5	7.0	3394	5	Q77384	Q77384 plasmodium
35	215.5	7.0	1395	2	Q9AISO	Q9AISO staphylococ
36	215.5	7.0	1564	2	Q931M3	Q931M3 staphylococ
37	213.5	6.9	2206	5	Q96205	Q96205 plasmodium
38	213	6.9	821	3	Q08581	Q08581 saccharomyc
39	212.5	6.9	1125	16	Q9BPM9	Q9BPM9 mycoplasma
40	212.5	6.9	3130	5	Q9BK46	Q9BK46 plasmodium
41	212	6.9	1946	5	Q97291	Q97291 plasmodium
42	212	6.9	2771	5	Q26216	Q26216 plasmodium
43	211.5	6.9	1344	2	Q49545	Q49545 mycoplasma
44	211.5	6.9	3724	5	Q77320	Q77320 plasmodium
45	211	6.8	691	5	Q9GSD1	Q9GSD1 plasmodium

ALIGNMENTS

RESULT	ID	PRELIMINARY:	PRT:	1092 AA.
070022	070022			
AC	070022			
DT	01-AUG-1998 (TREMblrel. 07, Created)			
DT	01-AUG-1998 (TREMblrel. 07, Last sequence update)			
DT	01-JUN-2001 (TREMblrel. 17, Last annotation update)			
DE	FIBRINOGEN-BINDING PROTEIN PRECURSOR.			
OS	Staphylococcus epidermidis.			
OC	Bacteria; Firmicutes; Bacillus/Clostridium group;			
CC	Bacillus/Staphylococcus group; Staphylococcus.			
OX	NCBI_TaxID=1282;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=HB:			
RX	MEDLINE=98261511; PubMed=9596732;			
RA	Nilsson M., Frykberg L., Flock J.I., Pel L., Lindberg M., Guss B.;			
RT	"A fibrinogen-binding protein of Staphylococcus epidermidis."			
RL	Infect. Immun. 66:2666-2673(1998).			
DR	EMBL: Y17116; CAA76638.1; -			
DR	InterPro: IPR001899; Gram_pos_anchor.			
DR	Pfam: PF00746; Gram_pos_anchor: 1.			
DR	PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.			
KW	Signal.			
FT	CHAIN			
FT	SIGNAL			
SO	SEQUENCE			
Query Match	97.7%	Score 3015;	DB 2;	Length 1092;
Best Local Similarity	99.7%;	Pred. NO. 6.2e-113;		
Matches 582;	Conservative	2;	Mismatches 0;	Indels 0;
DB	8 SDEKKNVNNOSINDDNNQIIKKETNNYDIERSERFESTNVDENFATFLQK 67			
OY				
DB	75 SDEKKNVNNOSINDDNNQIIKKETNNYDIERSERFESTNVDENFATFLQK 134			
OY				
DB	68 TPQDWTHTLEEFVKSSSVSSNSSIDTAQOPSHTTINRESVQTSADVEDSHVDFANS 127			
OY				

```

Db 135 TPQDNTHLTEEEVKSSSVSSSSSIDTAQOPSHHTINREESVQTSNDVEDSHVDFANS 194
Qy 128 KIKESNTESGKEENTIEOPNKKVEDSTTSQPSGYNIIDEKISNOBELNLPINEYENKAR 187
Db 195 KIKESNTESGKEENTIEOPNKKVEDSTTSQPSGYNIIDEKISNOBELNLPINEYENKAR 254
Qy 188 PLSTTSAOPSIKRVTVNOLAAEOGSNVNHLIKVTDOSTIEGYDSEGVYKADAEENLIYD 247
Db 255 PLSTTSAOPSIKRVTVNOLAAEOGSNVNHLIKVTDOSTIEGYDSEGVYKADAEENLIYD 314
Qy 248 VTFEYDDKVKSGDTMTVDIDKNTVPSDLTDSFTPIKIDNSGEIATGTVDKKNKQIYTT 307
Db 315 VTFEYDDKVKSGDTMTVDIDKNTVPSDLTDSFTPIKIDNSGEIATGTVDKKNKQIYTT 374
Qy 308 FTDYVDKXENIKAHKLKSYIDSKSVNNKTKLDEYKRTALSSVKKITVEYQRPENRNT 367
Db 375 FTDYVDKXENIKAHKLKSYIDSKSVNNKTKLDEYKRTALSSVKKITVEYQRPENRNT 434
Qy 368 ANLOSMFTNIDTKNHTVEOTIYINPLRSKAKETNNISGNGDEGSTIIDSTIIKVKYVG 427
Db 435 ANLOSMFTNIDTKNHTVEOTIYINPLRSKAKETNNISGNGDEGSTIIDSTIIKVKYVG 494
Qy 428 DNQNLPSNRIYDYSEEDVTNDYDQALGNNDVNIINGNDISPYIIKIVISKYDPNKDXY 487
Db 495 DNQNLPSNRIYDYSEEDVTNDYDQALGNNDVNIINGNDISPYIIKIVISKYDPNKDXY 554
Qy 488 TTIOQTVMTQITINEYTEGFRTASDYDNTIAFSTSSGOGGDLPPPKTYKIGDYVEDVDK 547
Db 555 TTIOQTVMTQITINEYTEGFRTASDYDNTIAFSTSSGOGGDLPPPKTYKIGDYVEDVDK 614
Qy 548 DGIQNTNDNEKPLSNVLTLYTPDGTSSKSVRTDEGKQYDPGVQ 591
Db 615 DGIQNTNDNEKPLSNVLTLYTPDGTSSKSVRTDEGKQYDPGVQ 658

RESULT 2
Q9KII3
ID 09KII3 PRELIMINARY: PRT: 931 AA.
AC 09KII3:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE FIBRINOGEN-BINDING PROTEIN SDRG.
GN SDRG.
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K28:
RX MEDLINE=20340957; PubMed=10878118;
RA McCrea K.W., Hartford O., Davis S., Ni Eidhin D., Lina G.,
RA Speziale P., Foster T.J., Hook M.;
RT "The serine-aspartate repeat (Sdr) protein family in Staphylococcus
RT epidermidis."
RL Microbiology 146:1535-1546(2000).
DR EMBL: AF245042; AAF72510.1;
DR InterPro: IPR001899; Gram_pos_anchor.
DR Pfam: PF00746; Gram_pos_anchor.1;
DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN.1.
SQ SEQUENCE 931 AA; 102955 MW; 591E657D97027116 CRC64;

```

```

Query Match 91.1%; Score 2812.5; DB 2; Length 931;
Best Local Similarity 93.3%; Pred. No. 6.2e-105;
Matches 545; Conservative 17; Mismatches 21; Indels 1; Gaps 1;
Qy 8 SDEKKNVNNOSINDNNQIIKKEETNNYDIEKRSERSTSTNNVDENAEFTLOK 67
Db 74 SSNEKNNDVNNOSINDDDNQ-IKKEETNSNDIAENNSKDIYQSTTNVDENAEFTLOK 132
Qy 68 TPQDNTHLTEEEVKSSSVSSSSSIDTAQOPSHHTINREESVQTSNDVEDSHVDFANS 127

```

```

Db 133 TPQDNTHLTEEEVKSSSVSSSSSIDTAQOPSHHTINREESVQTSNDVEDSHVDFANS 192
Qy 128 KIKESNTESGKEENTIEOPNKKVEDSTTSQPSGYNIIDEKISNOBELNLPINEYENKAR 187
Db 193 KIKESNTESGKEENTIEOPNKKVEDSTTSQPSGYNIIDEKISNOBELNLPINEYENKAR 252
Qy 188 PLSTTSAOPSIKRVTVNOLAAEOGSNVNHLIKVTDOSTIEGYDSEGVYKADAEENLIYD 247
Db 255 PLSTTSAOPSIKRVTVNOLAAEOGSNVNHLIKVTDOSTIEGYDSEGVYKADAEENLIYD 312
Qy 248 VTFEYDDKVKSGDTMTVDIDKNTVPSDLTDSFTPIKIDNSGEIATGTVDKKNKQIYTT 307
Db 315 VTFEYDDKVKSGDTMTVDIDKNTVPSDLTDSFTPIKIDNSGEIATGTVDKKNKQIYTT 372
Qy 308 FTDYVDKXENIKAHKLKSYIDSKSVNNKTKLDEYKRTALSSVKKITVEYQRPENRNT 367
Db 375 FTDYVDKXENIKAHKLKSYIDSKSVNNKTKLDEYKRTALSSVKKITVEYQRPENRNT 432
Qy 368 ANLOSMFTNIDTKNHTVEOTIYINPLRSKAKETNNISGNGDEGSTIIDSTIIKVKYVG 427
Db 435 ANLOSMFTNIDTKNHTVEOTIYINPLRSKAKETNNISGNGDEGSTIIDSTIIKVKYVG 492
Qy 428 DNQNLPSNRIYDYSEEDVTNDYDQALGNNDVNIINGNDISPYIIKIVISKYDPNKDXY 487
Db 495 DNQNLPSNRIYDYSEEDVTNDYDQALGNNDVNIINGNDISPYIIKIVISKYDPNKDXY 552
Qy 488 TTIOQTVMTQITINEYTEGFRTASDYDNTIAFSTSSGOGGDLPPPKTYKIGDYVEDVDK 547
Db 555 TTIOQTVMTQITINEYTEGFRTASDYDNTIAFSTSSGOGGDLPPPKTYKIGDYVEDVDK 612
Qy 548 DGIQNTNDNEKPLSNVLTLYTPDGTSSKSVRTDEGKQYDPGVQ 591
Db 615 DGIQNTNDNEKPLSNVLTLYTPDGTSSKSVRTDEGKQYDPGVQ 656

RESULT 3
Q9KWX6
ID 09KWX6 PRELIMINARY: PRT: 1171 AA.
AC 09KWX6:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE BONE SIALOPROTEIN-BINDING PROTEIN.
GN BOP.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-024:
RX MEDLINE=20115096; PubMed=10642520;
RA Tung H.S., Guss B., Hellman U., Persson L., Rubin K., Ryden C.;
RT "A bone sialoprotein-binding protein from Staphylococcus aureus: a
RT member of the staphylococcal sdr family ]."
RL Biochem. J. 345:611-619(2000).
DR EMBL: Y18653; CAB75732.1;
DR InterPro: IPR001899; Gram_pos_anchor.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN.1.
SQ SEQUENCE 1171 AA; 127123 MW; C5BC812F9DA5A884 CRC64;

```

```

Query Match 37.7%; Score 1162.5; DB 2; Length 1171;
Best Local Similarity 42.8%; Pred. No. 4.4e-39;
Matches 271; Conservative 101; Mismatches 182; Indels 79; Gaps 20;
Qy 8 SDEKKNVNNOSINDNNQIIKKEETNNYDIEKRSERSTSTNNVDENAEFTLOK 58
Db 57 STENAKODEASASDNKEVSESTENSTQKNDLTPNKRKTVDSHQEKKEAPTSSSTQOQ 116
Qy 59 ENAEFTLOKTPQDNTHLTEEEVKSSSVSSSSSIDTAQOPSHHTINREESVQTSND 115

```

```
Db 117 ONNATSETKPN--IEKENVKPSTDKTATEDSVILLEKKAPANT--NNDVTKPS-- 170
Qy 116 VEDSHVSPANSKIKESNT---ESKEENTIEQPNKVED---STTSOPSGYTN--DEK 167
Db 171 -----TSEIOTPTPTQESTINENSOPQPPSKYDNOVDTTNEKEVNSKEEL 220
Qy 168 ISNOELNLPINE--YENKARPLSTSAQPSIKRV-----TVNOLAEOGSNNHLIK 219
Db 221 KNBPKELELVARNDSNTKSTKPVATAPTVAAPKRVNAKIRPAVAQPAVASNNVNDLT 280
Qy 220 VTPOSITGDESGEVIKAHAENLIYDVFEDDKVKSQDGMTVVDIDKNVPSDLTDSF 279
Db 281 VTKQITEGIND-DGVIQAHQHEHITYSDFEKIDNAKAGDMTKYKDETFIPSDITDFE 339
Qy 280 TTPKIKDSGEIIATGYDNKNKQITTYFTDYVDYKYNIAKHLKLSYIDSKVPPNNMTK 339
Db 340 TPVDITDSGGEVIYAKGTFDLNTKTYKFTDYVDYREVNNAKLEINSYIDKKEVP--NETN 398
Qy 340 LDVEKKTALSSVNTKITYEYORPNENRTANLQSMETNIDTKNHTVEQTIYINPLRSYAKE 399
Db 399 LNLTPATADKETSKNVKVEYOKPIYKDESNIQSIFSHLDTTQHEVEQTIYVPLKLNKN 458
Qy 400 TTNVNI-----SGNGD-----EGSTIIDSTIIIVKYKVDNONLPDENRIYDSEYEDVTND 450
Db 459 TTVTKSGGVADNGDYTGDSSTIIDSNTKIKYKVASQQLPQSNKTIYDQYEDVYNS 518
Qy 451 DYALQNN---NDVINFGNIDSPYIYKISKYDNKDDYTTIOQTVMOTITNEYTGEF 507
Db 519 --VTNKMYGTGMANINCGDIDSAIYKVKVSKYTPGAEEDDLAVOOGVRMTT-----TNKY 571
Qy 508 RTASVDN-----TAFSTSGOGGODLPEKTYKIGDYVWEDVDGQIONTNDNK 558
Db 572 NYSSAGTITLYQLTLTVVTVSVK-----PEEKLYKIGDYVWEDVDGQVGTDSKEK 626
Qy 559 PLSNVLVTLTYPDGTSKSYRTDEGDKYQFDGYO 591
Db 627 PMANVLVTLTYPDGTTKSVRTDANGHYERGLK 659

RESULT 4
086489 PRELIMINARY; PRT; 1166 AA.
AC 086489;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE SDR E PROTEIN.
GN SDR E.
OS Staphylococcus aureus.
Bacteria; Firmicutes; Bacillus/Clostridium group;
Bacillus/Staphylococcus group; Staphylococcus.
NCBI_TaxID=1280;
RA [1]
RN SEQUENCE FROM N.A.
RC STRAIN-NEWMAN.
RX MEDLINE=99098700; PubMed=9884231;
RA Josephson E., Mccrea K., NI Eldhim D., O'Connell D., Cox J., Hook M.,
RA Foster T.D.;
RT "Three new members of the serine-aspartate repeat protein multigene
RT family of Staphylococcus aureus.";
RL Microbiology 144:3387-3395(1998).
DR EMBL, AJ005647; CAA06652.1;
DR InterPro; IPR001899; Gram_Pos_anchor.
DR Pfam; PF00746; Gram_Pos_anchor.1.
DR ProSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN.1.
SQ SEQUENCE 1166 AA; 126548 MW; 750A7B0135287D4A CRC64;
```

```
Query Match 37.1%; Score 1146.5; DB 2; Length 1166;
Best Local Similarity 42.5%; Pred. No. 1.9e-38;
Matches 263; Conservative 110; Mismatches 191; Indels 55; Gaps 18;
```

```
Qy 8 SSDEEKNDVYINNNGSINTDNNQI--IKKEETNNYDGIKRSRDRPSTTNDVENDENATFL 65
```

```
Db 66 ATTSNKEVSETEENNSTTENNSTNPIKKE--TNTDSQPEAKKESTSSSTQKQNNVTAT 123
Qy 66 OKTPDNDTHLTREEYKES---SSVESSNSIDTQAQPSHTTINREESVQSDNVEDSHVS 122
Db 124 TETKPN--IEKENVKPSTDKTATEDSVILLEKKAPANT--NNDVTKPS-----TS 172
Qy 123 DEANSKIKESNT---ESKEENTIEQPNKVEDSTTSOPSGYTNDEKIS--NODELNLP 178
Db 173 EPSTSEIOTKPTPTQESTINENSOPQPPSKVD--NOVTATNTKEEVNSKEELKNP 229
Qy 179 -----INEYENKARPLSTSAQPSIKRV-----TVNOLAEOGSNNHLIKYDQ 223
Db 220 EKLKELVARNDSNTDSTKPVATAPTVAAPKRVNAKIRPAVAQPAVASNNVNDLTKYKQ 289
Qy 224 SITTEGYDSEGIYKAHAENLIYDVFEDDKVKSQDGMTVVDIDKNVPSDLTDSFTIPK 283
Db 290 TIKVG-DGKDNVAHAHDGDIETFEFTIDNKVKKGDMTINYDKNVIIPSDLTDRNDID 348
Qy 284 IKDSNGEIIATGYDNKNKQITTYFTDYVDYKYNIAKHLKLSYIDSKVPPNNMKLQVE 343
Db 349 ITDPGGEVIYAKGTFDKATKQITTYFTDYVDYKEDIKSRLLTYSYIDKKTVP--NETSLNLT 407
Qy 344 YKTALSSVNTKITYEYORPNENRTANLQSMETNIDTKNHTVEQTIYINPLRSYAKEVTNN 403
Db 408 FATACKETSQNTVYQDPMVHGDNSIQSIFTKLDEDKOTIEQIYVPLKRSATNTKYD 467
Qy 404 ISGNG-----DEGSTIIDSTIIIVKYKVDNONLPDENRIYDSEYEDVTND-DYA 453
Db 468 IAGSOVDYGNKILGNGSGTIIIDQNEIKYKVKVNSDQQLPQSNRIYDFQYEDVTSQFQNK 527
Qy 454 QLGNNNDVINFGNIDSPYIYKISKYDNKDDYTTIOQTVMOTITNEYTGEFTAYD 513
Db 528 KSESNVATFLDGDINSALTIKVKVSKYTPSDGELIDAGTSMTPT-DK1-GYVYAGIS 585
Qy 514 NTIAFSTSGOGGODL-PEKTYKIGDYVWEDVDGQIONTNDNKPLSNVLVTLTYPDG 572
Db 586 NPIVTSNDTGGDGYKPEEKIKYKIGDYVWEDVDGQVGTDSKRRPMANVLVTLTYPDG 645
Qy 573 TSKSVRTDEGDKYQFDGYO 591
Db 646 TTKSVRTDANGHYERGLK 664

RESULT 5
099W46 PRELIMINARY; PRT; 1141 AA.
AC 099W46;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE SER-ASP RICH FIBRINOGEN-BINDING, BONE STALOPROTEIN-BINDING
DE PROTEIN.
GN SDR OR SA0521.
OS Staphylococcus aureus (strain N315).
Bacteria; Firmicutes; Bacillus/Clostridium group;
Bacillus/Staphylococcus group; Staphylococcus.
NCBI_TaxID=15879;
RA [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.T., Nagai Y., Ito T., Kanamori M.,
RA Matsumaru H., Matuyama A., Murakami H., Hoshoyama A., Mizutani-U Y.,
RA Takahashi N.K., Sawano T., Inoue R.I., Kaito C., Sekimizu K.,
RA Hiraoka H., Kuhara S., Goto S., Yabuzaki J., Kanehisa M.,
RA Yamashita A., Oshima K., Furuya K., Yoshino C., Shibata T., Hattori M.,
RA Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
DR EMBL, AP003131; BAB41752.1;
DR InterPro; IPR001899; Gram_Pos_anchor.
```

DR Pfam: PF00746; Gram_pos_anchor.1.
 DR PROSITE: PS00343; GRAM_POS_ANCHORING. UNKNOWN.1.
 KW Complete proteome.
 SQ SEQUENCE 1141 AA; 124026 MW; 445419D0B8C5A4F8 CRC64;

Query Match 36.9%; Score 1139.5; DB 16; Length 1141;
 Best Local Similarity 43.1%; Pred. No. 3.6e-38;
 Matches 266; Conservative 99; Mismatches 193; Indels 59; Gaps 19;

QY 8 SDEEKNVYNNOSINTDNNQIIKKEETNNYDIEKRSDEESTSTNVDENEATFLQK 67
 DB 69 SDNKEVSETENNSTENDSTNPI--KKEETN-DSOPEAKEESTSTSSQOOQNNVTATTE 125
 QY 68 TPQDTHLLEEVKES---SSVESSNSSIDTAQPSHTTINREESVOTSDNVEDSHVDF 124
 DB 126 TKPQN--TEKENVKPSTOKTATEDSVILEEKKAPNT--NNDVTKPS----- 170
 QY 125 ANSKIKESNT---ESGKEENTIEOPNKYKEDSTTSOPSGYTNIDEKIS--NODELINLP-- 178
 DB 171 -TSEIQTPTPOESTNINENSOPQPTPSKVD---NQVTDATNPKRPVAVNSKEELKNNEBK 226
 QY 179 -----INEYENKARPIST--TSAQP-----SIRKVTYNQLAEGGSVNNHLIKYTDOSI 225
 DB 227 LKELYRNNDNDRSTKPAATAFTSVAPKRLNKKMFPAVAQPAVANSNNVNDLITVTKOTI 286
 QY 226 TEGYDSDSGVIAKHADEMLIYDVFPEVDKVKSGDTMTVDIDKNTVPSDLTDSFTIPKIK 285
 DB 287 KVG-DGKDNVAAAHDKGKIEVDTEFTIDNKVKKGDTMTINVDKNVIPSOLDIDKNDPIDIT 345
 QY 286 DNSGEIITGTIDNKKQITTYFTDYVDKYENIKAKHLTSTYIDSKVPPNNNTKLDVEYK 345
 DB 346 DPGSEVIAKGFTDKATKQITTYFTDYVDKYEDIKARLTYSYIDKQAVP--NETSLNTFA 404
 QY 346 TALSSVNTIYVEYORPENRNTANOSMFTNIDTNHVEQITTYINPLRYSAKKEINVIS 405
 DB 405 TAGKETSQNVSVYDQDPVHSGSNISQIFTKLIDENKQITEQIYVNPFLKKTATNKKVDA 464
 QY 406 GNG-----DEGSTIIDSTIIKVKYKVGDNQMLPDSNRIYDSEYEDVTND--DYAQL 455
 DB 465 GSQVVDYGNIKLGNSTIIDQNTTEIKVKYKVNQOQLPQSNRIYDSEYEDVTSQDNKKS 524
 QY 456 GNNNDVNINFGNIDSPYIIKVIKSKYDPNKDDYTTIQQVTVMTQTTINETGEERTASYDNT 515
 DB 525 FSNNAVATLDFGIDNSAVYIIKVIKSKYTPSDGLDIAQGSMTT--DKY-GYNYVAGYSNF 582
 QY 516 IAFSTSSGQGGDL--PPEKTYKIGDYVWEDVDKGIQNTNDEKPLSNVLTLYTPPDGS 574
 DB 583 IYTSNDTGGGDTVPPEKLYKIGDYVWEDVDKDVQGSDEKEMANVLTLYTPPDGTT 642
 QY 575 KSVRTDEGKYQFDGVQ 591
 DB 643 KSVRTDANGHYEFGGLK 659
 RESULT 6
 Q932F7 PRELIMINARY; PRT; 1141 AA.
 ID Q932F7;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE SER-ASP RICH FIBRINOGEN-BINDING, BONE STALOPOTEIN-BINDING
 DE PROTEIN.
 DE SDR OR SAV0563.
 OS Staphylococcus aureus (strain Mu50).
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Staphylococcus.
 OX NCBI_TaxID=158878;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21311952; PubMed=11418146;
 RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,

RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iian J.-Q., Ito T.,
 RA Kanamori M., Matsunaru H., Maruyama A., Murakami H., Hosoyama A.,
 RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
 RA Sekimizu K., Hiraoka H., Kubara S., Goto S., Yabuzaki J.,
 RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
 RA Hattori M., Ogasawara N., Hayashi H., Hiratazu K.,
 RT "Whole genome sequencing of methicillin-resistant Staphylococcus
 RT aureus.";
 RL Lancet 357:1225-1240(2001).
 DR EMBL: AP003359; BAB56725.1; .
 KW Complete proteome.
 SQ SEQUENCE 1141 AA; 124038 MW; E679FC2991846D9 CRC64;

Query Match 36.9%; Score 1139.5; DB 16; Length 1141;
 Best Local Similarity 43.1%; Pred. No. 3.6e-38;
 Matches 266; Conservative 99; Mismatches 193; Indels 59; Gaps 19;

QY 8 SDEEKNVYNNOSINTDNNQIIKKEETNNYDIEKRSDEESTSTNVDENEATFLQK 67
 DB 69 SDNKEVSETENNSTENDSTNPI--KKEETN-DSOPEAKEESTSTSSQOOQNNVTATTE 125
 QY 68 TPQDTHLLEEVKES---SSVESSNSSIDTAQPSHTTINREESVOTSDNVEDSHVDF 124
 DB 126 TKPQN--TEKENVKPSTOKTATEDSVILEEKKAPNT--NNDVTKPS----- 170
 QY 125 ANSKIKESNT---ESGKEENTIEOPNKYKEDSTTSOPSGYTNIDEKIS--NODELINLP-- 178
 DB 171 -TSEIQTPTPOESTNINENSOPQPTPSKVD---NQVTDATNPKRPVAVNSKEELKNNEBK 226
 QY 179 -----INEYENKARPIST--TSAQP-----SIRKVTYNQLAEGGSVNNHLIKYTDOSI 225
 DB 227 LKELYRNNDNDRSTKPAATAFTSVAPKRLNKKMFPAVAQPAVANSNNVNDLITVTKOTI 286
 QY 226 TEGYDSDSGVIAKHADEMLIYDVFPEVDKVKSGDTMTVDIDKNTVPSDLTDSFTIPKIK 285
 DB 287 KVG-DGKDNVAAAHDKGKIEVDTEFTIDNKVKKGDTMTINVDKNVIPSOLDIDKNDPIDIT 345
 QY 286 DNSGEIITGTIDNKKQITTYFTDYVDKYENIKAKHLTSTYIDSKVPPNNNTKLDVEYK 345
 DB 346 DPGSEVIAKGFTDKATKQITTYFTDYVDKYEDIKARLTYSYIDKQAVP--NETSLNTFA 404
 QY 346 TALSSVNTIYVEYORPENRNTANOSMFTNIDTNHVEQITTYINPLRYSAKKEINVIS 405
 DB 405 TAGKETSQNVSVYDQDPVHSGSNISQIFTKLIDENKQITEQIYVNPFLKKTATNKKVDA 464
 QY 406 GNG-----DEGSTIIDSTIIKVKYKVGDNQMLPDSNRIYDSEYEDVTND--DYAQL 455
 DB 465 GSQVVDYGNIKLGNSTIIDQNTTEIKVKYKVNQOQLPQSNRIYDSEYEDVTSQDNKKS 524
 QY 456 GNNNDVNINFGNIDSPYIIKVIKSKYDPNKDDYTTIQQVTVMTQTTINETGEERTASYDNT 515
 DB 525 FSNNAVATLDFGIDNSAVYIIKVIKSKYTPSDGLDIAQGSMTT--DKY-GYNYVAGYSNF 582
 QY 516 IAFSTSSGQGGDL--PPEKTYKIGDYVWEDVDKGIQNTNDEKPLSNVLTLYTPPDGS 574
 DB 583 IYTSNDTGGGDTVPPEKLYKIGDYVWEDVDKDVQGSDEKEMANVLTLYTPPDGTT 642
 QY 575 KSVRTDEGKYQFDGVQ 591
 DB 643 KSVRTDANGHYEFGGLK 659
 RESULT 7
 O86488 PRELIMINARY; PRT; 1315 AA.
 ID O86488
 AC O86488;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE SDRD PROTEIN.
 GN SDRD.
 OS Staphylococcus aureus.

Db 553 NALGFTNNOSGAG-----GEVYKIGYVWEDTRKNGOEL--GEKGVANTYTV-PDNN 605

QY 574 SKSVR--TDEGKY 585

Db 606 NTKVGEAVTKEDGSY 620

RESULT 9

Q9K14 PRELIMINARY: PRT: 1733 AA.

AC 09K14

DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

DE PUTATIVE CELL-SURFACE ADHESIN SDRF.

GN SDRF.

OS Staphylococcus epidermidis.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillus/Staphylococcus group; Staphylococcus.

NCBI_TaxID=1282;

SEQUENCE FROM N.A.

RC STRAIN-9491;

RX MEDLINE-20340957; PubMed-10878118;

RA McCrea K.W., Hartford O., Davis S., Nl Eichlin D., Lina G.,

RA Speziale P., Foster T.J., Hook M.,

RT "The serine-aspartate repeat (Sdr) protein family in Staphylococcus

epidermidis."

RT Microbiology 146:1535-1546(2000).

RL EMBL: AF245041; AAF72509.1; -

DR InterPro: IPR001899; Gram_pos_anchor.

DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.

SQ SEQUENCE 1733 AA; 184720 MW; D8D62EAL692FDE8 CRC64;

Query Match 16.4%; Score 506; DB 2; Length 1733;

Best Local Similarity 27.4%; Pred. No. 9, 6e-13;

Matches 174; Conservative 101; Mismatches 276; Indels 84; Gaps 27;

QY 7 PSSDEE---KNDYINNOSINTDNNQIIKEETNNYDIKESREDTST----- 54

Db 135 PTVNESISAEPTKSTTODSTTEKNPSTL-KDNLSSSTSKSDESHTKQAQSTNK 193

QY 55 TANDENEA-TFLQKT-PQONTHTLEEVKESSESSVSNS-----SIDTQAQPSHTT 103

Db 194 SNLDITDPSQSEKTSQAANDSTNOASAPSKQDSKPEQKYKTKFDEPTQDVEHTT 253

QY 104 IN-REESVQTSQNVEDSHVSDFAANSKIKESNTESGKEENTIEQPNKVKEDSTTSQPSGYT 162

Db 254 TKLKTPTSVSTDSVNDK--QDYTRSAV---ASLGVDSMTEFATNVAQNDNLKASRE 307

QY 163 NIDEKI-----SNODELMLP--INYEKARPSTLSAOPSIKRYVNOQLAE-- 209

Db 308 QINEAIIAEALKKDFNSNPYGVDTPLALNRSQSKNSP--HKSASP--RNLMLSLAEEN 362

QY 210 QGSNVAHLIKVQDO--SITEGYDSEGVYKAHDAENLIYDVEYDQVSKSGTMTVDID 267

Db 363 SGANVADKXKITPTLSLKNNSNHANNVIMPTSEQPNLKANELEDISKEGTFPIKKG 422

QY 268 KNTVPSDLTDSFTPIKIDNSGEIATGTYNKKNQITVTFDIDYDKENIKAKHLTSTY 327

Db 423 QYRPGGLELPAIKTQLRSKDSIVANGVYDKTTNTTTFYTVYDQYQVITGSFLIAT 482

QY 328 IDSKSVNNNTKIDVEYKATLASVKNKTIYVEYRPENNTANLQSFNTIDTAKNHYEOT 387

Db 483 PKRETAIKQNQYPMETJIANEVKKDFIYDGNKKDNTT--AAVANVDNNKHNHY 539

QY 388 IYVNPRLYSAKETNVNINSGDEGTTIIDSTI--IKYKVGDNQNLPS--NRIDYSE 443

Db 540 VYINQNNQNPKYAKY-----FSTVKNGEFIPGEVKKYIEVDTNAAMDSEFPLDINSSN 591

QY 444 YEDVTNDVYAOI-GNNNDVINFGNI---DSPYIIKIVISKYDPNKKDYTTIQQVTMTQT 499

Db 592 VKDVTQSPAPKXSADSTRDINFAFSMANGKRYIYQAVRPTGTGNYT--EYWLTRDGT 649

QY 500 INEYTGEPRTASYDNTIAFTSTSSGQGGDLPEPKTYIKGIDYVWEDVXDGIQNTDNEKP 559

Db 650 TN--TNDYRGKSTTVYTLNCSSTQAQGNP---FTSLDGYWLDKNNKGYO--DDDEKG 702

QY 560 LSNVLYTLTPPGTSGK---VRTDEGKYQFQGVQ 591

Db 703 LAGVYVTL--KDSNNRELQRYTQDSGHYQFDNLQ 735

RESULT 10

Q99W48 PRELIMINARY: PRT: 953 AA.

AC 099W48

DT 01-JUN-2001 (TREMBLrel. 17, Created)

DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE SER-ASP RICH FIBRINOGEN-BINDING, BONE STALOPROTEIN-BINDING

DE PROTEIN.

GN SDRF OR SA0519 OR SAV0561.

OS Staphylococcus aureus (strain N315), and

OS Staphylococcus aureus (strain Mu50).

OC Bacteria; Firmicutes; Bacillus/Clostridium group;

OC Bacillus/Staphylococcus group; Staphylococcus.

OK NCBI_TaxID=158879, 158878;

OX

RM

RP SEQUENCE FROM N.A.

RC SPECIES=S.aureus (strain N315), and S.aureus (strain Mu50);

RX MEDLINE-21311952; PubMed-11416146;

RA Kurda M., Ohta T., Uchiyama I., Baba T., Yizawa H., Kobayashi I.,

RA Cul L., Oguchi A., Aoki K., I., Nagai Y., Lien J.-Q., Ito T.,

RA Kanemori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,

RA Mizutani U.Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,

RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,

RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,

RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.,

RT "Whole genome sequencing of methicillin-resistant Staphylococcus

aureus."

RT Lancet 357:1225-1240(2001).

RL EMBL: AP003131; BAB41750.1; -

DR EMBL: AP003359; BAB56723.1; -

DR InterPro: IPR001899; Gram_pos_anchor.

DR Pfam: PF00746; Gram_pos_anchor_1.

DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.

KW Complete proteome.

SQ SEQUENCE 953 AA; 103292 MW; 729A7169A074A1E5 CRC64;

Query Match 15.9%; Score 490; DB 16; Length 953;

Best Local Similarity 28.2%; Pred. No. 2, 3e-12;

Matches 162; Conservative 83; Mismatches 236; Indels 94; Gaps 22;

QY 42 GIEKRSDETESTTNDENEATFLQKTPDONTHTLEEVKESSESSVSNSIDTQAQPSH 101

Db 45 GHEAKAAEHTNENLQSKNETT---APSENK--TTEKV--DSROLKDNVQYATAPQAPV 96

QY 102 TTNREESVQTSQNVEDSHVSDFAANSKIKESNTESGKEENTIEQPNKVKEDSTTSQPSGY 161

Db 97 T-----MSDSATVKETSSNNQS-----PONATASOSTTQTSNV 129

QY 162 TNDIEK-----ISNODELMLPINEYEN-KARPLSTSAOPSIKRYVNOQLAE--EAGSNVNH 216

Db 130 TTNDSSTTYSNETDKSNL--TQAKNVSTTPKTTITKQRLNMAAVNTAAAPQAGTNVD 187

QY 217 LKIVTDSI-----TEGYDSEGVYKAHDAENLIYDVEYDQVSKSGDTMT 263

Db 188 KYHFTNIDLAIDKGNKTKTQNTFEWATSVDLK-----LKANYTIDDSVKEDDTFT 239

QY 264 VDIKNTVPSDLTDSFTPIKIDNSGEIATGTYNKKNQITVTFDIDYDKENIKAKHL 323

Db 240 FRYGQYFRFGSVRLPQOTONLYNAQGNIIAKGIYDSKTWTNTTTFYTVYDQYQVITWVSGSFE 299

324 LRSYIDSKVNNNTKLDVEYKTAALSSVNTTIVEYORPNEBNTANLQSMFTNIDTKNH 383
 Db 300 QVAFKREKATTDKTAKEVTLGNDYTSKDVYD---GNOKQOOLISSTNYINNEDLS 356
 QY 384 VEOTIYIN-PLKRSKATNY-NISGNDEGSTIIDSTIIKVKYGVGNOLPDSNRLEYD 441
 Db 357 RNNIVYVNOPEKRYTKEFTETNLT-----GKENPDKAKNKIYEDVQNGFVDS-FTPDT 410
 QY 442 SEVEDYND-DYQOLGNNNDVINFGN----IDSPYIKVSKYDPNPKDDYTTIQGTVM 496
 Db 411 SKYKDVTCGQDVYISNDKATVDLNGOSSDQYIIQVAYADNSDNGKIDYLET 470
 QY 497 QTTINNEYGEFRASDYNTIAFTSSGOGGDLPEPEKTYKIGDYVWEVDKDIQNTNDN 556
 Db 471 QNGKSSMSN-----SYSVNGSSTAND-----QKKYNGDYVWEDTNDKQDA--N 516
 QY 557 EKPLSNLVLTITPDGTS-KSVRTDEDECKYQFDG 590
 Db 517 EKGIKGVYVILKDSNGKELDRTTIDENKGYQFTGL 551
 SUIT 11
 3MH7 PRELIMINARY: PRT; 881 AA.
 AC 093MH7;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE RECEPTIN FBL PRECURSOR.
 GN FBL.
 OS Staphylococcus lugdunensis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 CC Bacillus/Staphylococcus group; staphylococcus.
 OX NCBI_TaxID=28035;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-2342;
 RA "A fibronogen-binding protein of Staphylococcus lugdunensis."
 RT Submitted (Aug-2001) to the EMBL/Genbank/DBJ databases.
 RL EMBL; AF040823; AAK95649.1;
 DR EMBL Signal.
 KM SIGNAL.
 FT CHAIN 1 39 POTENTIAL.
 FT SIGNAL 40 881 RECEPTIN FBL.
 SQ SEQUENCE 881 AA; 94251 MW; D4296C4959CAFL9B CRC64;

Query Match 14.9%; Score 460.5; DB 2; Length 881;
 Best Local Similarity 27.3%; Pred. No. 3.1e-11;
 Matches 154; Conservative 88; Mismatches 231; Indels 91; Gaps 23;

13 KNDVINNNOSINTDNNQIIRKEETNNYDQIEKRSDEGSTINVDENATFLQK--TPQ 70
 Db 48 KQOIQHNNNA--TGGDT-----ODNNYN-----NEISQETTKQKOTLOS 86
 QY 71 DNTHLFEEEKSESSVESSNSIDTAQOPSHTTINREESVQTSNDYEDSHVSPFANSK 130
 Db 87 DNANSEAOAINEISD-----SHRYVTK--ATEALDNNSTLMTSTVSPATK 130
 QY 131 ESNTESGK--EEN--TIEOPNKVKEDSTTSQPSGYTNIDEKISN--DELLNLPINEYE 183
 Db 131 QDPTTISQGTQENNDATQTKTKYKODGNNNVLSQVATNDNOSNCPNSHLNTSTVTYN 190
 QY 184 N--KARPLSTTSQAOPSKRYTVNOLAEOGSNVNHLIKVDOSTIEGYDSEGYKAKDA 241
 Db 191 NNHVOBLAVYEATNTNNNTQTSIDISNKLSTVATTEAD-----TTPYHKA 238
 QY 242 ENLIVDTFEVDKVSAGDFTVDIDKRTVPSDLTDSFTTPIKIDNGSELIATGYDNK 301
 Db 239 EYVNLNRFQAPDVOAGDSIKITIPQALNNGYATAKAPINMAGD--QILATGIDBEG 297
 QY 302 KQITFTFDVYDKYKAHLKLTSYIDSKSVYNNNTKLDVEYKTAALSSVNTTIVEYOR 361

298 NLI-YFTFDVFNKNNITGOISIPGYIDPKNVHTG-KVNLFTSIGOTAKKVTYDYK 355
 QY 362 PNEKRTANLOSMFTNIDTKHTEQTIYINPLRSKATNYN--ISGN--GDEGSTIID 416
 Db 356 YGEFRNLISGTTIDQIDKRVNNTYKQIVYVNP-----SSDYVDPLRGGSIPGTSNVIID 411
 QY 417 D-STPIKVKYKGNQNDPDSNRIVYSEYEDYNTDDYAOGLNNNDVINFG--NIDSP 471
 Db 412 EONTSIVYKVKYKAHLTDSYV--DPSNEDVTSVKITEPEKGIYQINFTEDDQIISP 470
 QY 472 YIKIVISKYDPNKKDDYTTIQGTVMQTTINEYGEFRAS--YNTIAFTSSGOGGDL 529
 Db 471 YVVVINGHVDPNSNG-----NLVLRSTLYGDSNFTVRSAMUNVEYTHAGSNGGID 524
 QY 530 PPEKTYKID-----YVWEDYKD 548
 Db 525 KVPIDQGDSDAYSDSDADSD 548

RESULT 12
 086487
 ID 086487 PRELIMINARY: PRT; 947 AA.
 AC 086487;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE SDRK PROTEIN.
 GN SDRK.
 OS Staphylococcus aureus.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 CC Bacillus/Staphylococcus group; staphylococcus.
 OX NCBI_TaxID=1280;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-NEWMAN;
 RX MEDLINE=99098700; PubMed=9884231;
 RA Josefsson E., McCrea K., M Eklund D., O'Connell D., Cox J., Hook M., Foster T.J.;
 RA "Three new members of the serine-aspartate repeat protein multigene family of Staphylococcus aureus."
 RT Microbiology 144:3387-3395(1998).
 RL EMBL; AJ005645; CAA06650.1;
 DR EMBL; InterPro: IPR001899; Gram_pos_anchor.
 DR pfam: PF00746; Gram_pos_anchor; 1.
 DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN 1.
 SQ SEQUENCE 947 AA; 10288 MW; 3C6EFD6E35121554 CRC64;

Query Match 14.8%; Score 456; DB 2; Length 947;
 Best Local Similarity 27.1%; Pred. No. 5.1e-11;
 Matches 158; Conservative 94; Mismatches 222; Indels 108; Gaps 25;

42 GIEKRSDETESTNVDENATFLQKTPDNNHLEBEVKESSEVS-----SSSIDTAQ 97
 Db 45 GHEAKAHEHTNGELNOSKNETT-----APSEN-----KTKRVYSROLKDTQATATD 92
 QY 98 QPSHTTINREESVQTSNDVEDSHVSPFANSKIKESNTESGKEENTIQPKVKEDSTTSQ 157
 Db 93 QPKVT-----MSDSATVKESSNNQS-----PONATANGSTK 125
 QY 158 PSGYTNIDEK--ISNDELNLPINEYENKARPLSTTSQAOPSK-----RYTVNOLA- 208
 Db 126 TSNVTTNDKSSSTYSNETDSNL-----YQAKVSTTPPTTTIKPTLRBMVAVNYAAP 179
 QY 209 EOGSNVHLIKVT--DOSITEGY-DSEGYK--AHDAENLIDVTEVDKYSAGGTMT 263
 Db 180 QQGNVNDKVAHFSIDLAIDKGVNQTGKTEFWATSSDVLKANKYTIIDSVYEGDFT 239
 QY 264 VDIKNTVPSDLTDSFTTPIKIDNGSELIANGYDNKNKQTYFTFDYDYEYKAKHLK 323
 Db 240 FKYGQYFRPGSVRLPSQOTNNYNAGNITAKGIDSTNTTITTYFTNRYVDQYTVNRSGFE 299
 QY 324 LRSYIDSKVNNNTKLDVEYKTAALSSVNTTIVEYORPNEBNTANLQSMFTNIDTKNH 361

```

Db 300 QVAFARAKNATDTKTAAYKKEVTLGNDTYSEILVDY---GNKRAQBLISTSTNYINNEDLS 356
Oy 384 VEDTIIN-PLARSARETNV-NISGNGDESGTIIIDSTIIKVKVDONOL-----PDSN 436
Db 357 RNMATAVNOPKNTYTKOTFTYNTL-----GYKRNPAKNEKIYEVDQONFVDSFPDTS 411
Oy 437 RIVYSEYEDVTDNDYAQALGNNDVNIINGNIDS--PYIIKVI-----SKYDPNKDDYTT 489
Db 412 KIKVDYDQFVY---YSNDKKTATVDLMKGQSTSSNKQYIIQVAVPDSNDSTNGKIDY-- 466
Oy 490 IQQVMTQITINYEGEFPTASTDNTIATSTSGGCGGLDPEKTKIGDYVWEDVDKG 549
Db 467 -----TLDDDKTKYSW---SNSYSNVNGSSTANGD-----QKRYNLGDYVWEDTNKDG 511
Oy 550 IONTNNEKPLSNVLTATYTPDGTG--KSVATEDGKYPDGV 590
Db 512 KODA--NEKIGKVYILKDSNGEKELDRITTDENKGTQFTGL 551

```

```

RESULT 13
Oy 053653 PRELIMINARY: PRT: 933 AA.
AC 053653;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE CLUMPING FACTOR.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Staphylococcus.
OX NCBI_Taxid=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NEUMAN;
RX MEDLINE=94224142; PubMed=8170386;
RA McDevitt D., Francois P., Vaudaux P., Foster T.J.;
RT "Molecular characterization of the clumping factor (fibrinogen receptor.
RT of Staphylococcus aureus."
RL MOL. Microbiol. 11:237-248(1994).
DR EMBL: Z18852; CAA79304.1;
DR InterPro: IPR000515; BPD.transp.
DR InterPro: IPR001899; Gram_pos_anchor.
DR PROSITE: PS00402; BPD_TRANSP_INN_MEMBER; UNKNOWNM.1.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWNM.1.
SQ SEQUENCE 933 AA; 97058 MW; EB51A6DE2FE759F4 CRC64;

```

Query Match 14.6%; Score 452; DB 2; Length 933;
 Best Local Similarity 26.6%; Pred. No. 7.3e-11;
 Matches 154; Conservative 105; Mismatches 234; Indels 86; Gaps 26;

```

Oy 8 SDEEKNDVYINNOSINTDDNNQIIKKEETNNNDGIEKRESDRESTYNDENATLQK 67
Db 51 SNEKSNDSVSAAPKTDNTN-----SDTKSSNTN---NGETISVAQ 91
Oy 68 TPQDNHILTEEVKESVSSNSIDTAQOPSHTTINRESVQTSNVEDSHVSDPANS 127
Db 92 NP-----AQOETQSSSTNATTEETPVGTGATTTTNOANTPATQOS-SNTNAEELVN- 143
Oy 128 KIKESNTSGKEENTIQPNKVKEDSTTSQSGYTINDEKISNDELIN-LPINEYENKA 186
Db 144 --QTSNETTNDNTV-----SSVNSPQNSTNAENSTQCDTSTEAFTPSN---NES 189
Oy 187 RPLSTSAOPSIRKVTYN-----QLAAEQGSNNHLIKVTDQ--SITEGYDSEGV 235
Db 190 APOSTDSNKNQVNOAVNTSAPRRARAFSLAADAAPAGDTITNOALNTVGI-DSGTT 248
Oy 236 IKAHDAENLIYDTFEVDKVKSGDTMTVIDIKNTVPSDLTSPFIKIKDNGSEIATG 295
Db 249 VYRQAGVYKNTNGFSVPNSAVKQDFTKITVPRKELNLNGVSTAKVPPPIAAGD-QVLANG 307
Oy 296 TYDNKKNOITTFYDYDKENIKAHKLITSYIDKSKVPPNNNTKL-DVEIKTALSS--VN 352

```

```

Db 308 VIDS--DGNVIYFTFDVNRKRDVKATLTPAYID-----PENVKKTGNVTLATIGSTIAN 362
Oy 353 KRTVEYQRPENRRTANLOSMEFTNIDTKNHTVEQTIYINPL--RYSARETNVINSNGDE 410
Db 363 KTVLVDEYKGRFYNLSIGTIDQIDKTNNTYQTIYVPSGQVTAAPVLGRLKRNQTS 422
Oy 411 GSTIIDSSTIIKVKVGNQNLDPDSNRITYDYSEYEDVTNDYQAOLGNNDVNIING- 466
Db 423 NALIDQONTSIKYKVDNADLSESYFV--NPENFEDVTSVNTTFPNPQYKVEFNTPPD 481
Oy 467 NIDSPYIIKYSKYPDN-KDYTTIQOYTMQNTINEYGE--FRASVDNTIATSTSG 523
Db 482 QITTPYIVVNGHIDNSKGD-----LALRSTLGYNSNIITWRMSMDNEVAFFNNGSG 534
Oy 524 OGOG-DLP--PEKTKIGDY--VWEDVDKQ-GIONTNDN 556
Db 535 SGDGIKPPVPEQDPDELEPIPEDSDSDPSGSDSGS 573

```

```

RESULT 14
Oy 0932C5 PRELIMINARY: PRT: 935 AA.
AC 0932C5;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE FIBRINOGEN-BINDING PROTEIN.
GN FNB OR SAV0811.
OS Staphylococcus aureus (strain Mu50).
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Staphylococcus.
OX NCBI_Taxid=158878;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21111952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iino J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Uji Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus."
RL Lancet 357:1225-1240(2001).
DR EMBL: AP003360; BAB56973.1;
KW Complete proteome.
SQ SEQUENCE 935 AA; 96950 MW; DC5A2D92CE3BA91C CRC64;

```

Query Match 14.6%; Score 452; DB 16; Length 935;
 Best Local Similarity 25.9%; Pred. No. 7.3e-11;
 Matches 152; Conservative 109; Mismatches 224; Indels 102; Gaps 27;

```

Oy 8 SDEEKNDVYINNOSINTDDNNQIIKKEETNNNDGIEKRESDRESTYNDENATLQK 67
Db 51 SNEKSNDSVSAAPKTDNTN-----SDTKSSNTN---NGETISVAQ 91
Oy 68 TPQDNHILTEEVKESVSSNSIDTAQOPSHTTINRESVQTSNVEDSHVSDPANS 127
Db 92 NP-----AQOETQSSSTNATTEETPVGTGATTTTNOANTPATQOS-SNTNAEELVN- 143
Oy 128 KIKESNTSGKEENTIQPNKVKEDSTTSQSGYTINDEKISNDELIN-LPINEYENKA 186
Db 144 --QTSNETTNDNTV-----SSVNSPQNSTNAENSTQCDTSTEAFTPSN---NES 189
Oy 187 RPLSTSA-----OPSIKRVTVNOAA-----EQGSNNHLIKVTDQSTTEGYDD 231
Db 190 APOSTDSNKNQVNOAVNPSTPRARAFSLAADAAPAGDTITN--QLTDVKYKT---ID 244
Oy 232 SEGVIKHAHDAENLIYDTFEVDKVKSGDTMTVIDIKNTVPSDLTSPFIKIKDNGSEI 291

```

```

Db 245 SGTIVYPHOAGYKLVNGFSPNSAVKGDFTKIVPEKELNNGVTSTAKVPPIMAGD-QV 303
QY 292 IATGYDNKKKQIHYFTDVKDKENIKAKHLKLSYIDSKVPNNKTL-DVEYKTL-- 348
Db 304 LANGVIDS-DGNVYTFDVKDKENKENTANITMPAYID----PENVTKGAVTLTGIGT 358
QY 349 SSVNKTIVVEYQRENNERTANLQSMFTNIDTKNHTVEQTIYINPLRSKATENV----- 402
Db 359 NTSKTYLIDYEKYGQPHNLSIKGTIDQIDKTNNTYQTIYVNE----SGDNVVLPAIT 413
QY 403 -NISNGDESGTIIDSTIIKVKYVGNQNLPSNRIYDYSEYEDVTNDYAOIGNNNDV 461
Db 414 GNLIPTKSNALIDAKNTDIKRYRV-DNANDLSESYVNSDFEDVTNQVRISFPNANQY 472
QY 462 NINF---GNIDSPYIIKVIKSKYDPNKKDYTTTQQVTMQTTINEYTGEE--RTASYDNT 515
Db 473 KVEFPDDDDQITTPYIVVNGHIDPASTG-----DLALSTFYGYDSNFIWRSMSWDNE 526
QY 516 IASTSSGOGG-DLP--PEKTYKIGDY--VWEDVDKD-GIONTNDN 556
Db 527 VAFNNGSGDGDIDKPVVPEQPEGEIEPIPEDSDSPSGSDS 573

RESULT 15
Q99VJ4 PRELIMINARY; PRT; 989 AA.
AC Q99VJ4;
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE FIBRINOGEN-BINDING PROTEIN A, CLUMPING FACTOR.
GN CLFA OR SA0742.
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Staphylococcus.
OX NCBI_TaxID=158879;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.I., Nagai Y., Iino T., Kanamori M.,
RA Matsumaru H., Maruyama A., Murakami H., Hosoyama A., Mizutani-Ui Y.,
RA Takahashi N.K., Sawano T., Inoue R.I., Kaito C., Sekimizu K.,
RA Hiraoka H., Kuhara S., Goto S., Yabuzaki J., Kanehisa M.,
RA Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T., Hattori M.,
RA Ogasawara N., Hayashi H., Hiratsuku K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus."
RL Lancet 357:1225-1240(2001);
EMBL: AP003131; BAB41975.1;
InterPro: IPR000515; BPD_transp.
InterPro: IPR001899; Gram_pos_anchor.
PROSITE: PS00402; BPD_TRANSP_INN_MEMBER, UNKNOWN_1.
KW PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
SQ Complete proteome.
SQ SEQUENCE 989 AA; 102407 MW; DA6E807539623467 CRC64;

```

Query Match 14.6%; Score 452; DB 16; Length 989;
 Best Local Similarity 25.9%; Pred. No. 7.7e-11;
 Matches 152; Conservative 109; Mismatches 224; Indels 102; Gaps 27;

```

QY 8 SSDEKNDVYNNQSIINTDNNQIIEKEETNNYDGIKRSERDRTSTNTVDENATPLOG 67
Db 51 SNEKSNDSVSAAPRTDFTN-----SDTKSSNTN--NGETISVAQ 91
QY 68 TPQDNTHLIEEYKSSVSNSSIDTAAQPSHTTINREESVQTSNVEDSHVDFANS 127
Db 92 NP-----AQOETQSSSTNATTEEPVTGEATTTTNOANTPATQS-SNTNAEELVN- 143
QY 128 KIKESNTSGKEKENTIQPKVKEDSTTSQSGYTNIDEKISNDEJLN-LPINEYENKA 166
Db 144 --QTSNETSNDNTNIV-----SSVNSPQNSTNANTNVTQDTNSTEATPSN--NES 189

```

```

QY 187 RPLSTTSA-----QPSIKRVTVNQLAA-----EQQSNVHLIKVTDQSTTEGYD 231
Db 190 APQNTDASNNKDVYSAVNPSTPRHRAFSLAVAADAPAGTDITN--QLTDVAKYT---ID 244
QY 232 SEGVIKADHAEENLIYDVFEEVDKVKSGDTMTVDIDKNVPSSDITDSFTIPKIDNSGEI 291
Db 245 SGTIVYPHOAGYKLVNGFSPNSAVKGDFTKIVPEKELNNGVTSTAKVPPIMAGD-QV 303
QY 292 IATGYDNKKKQIHYFTDVKDKENIKAKHLKLSYIDSKVPNNKTL-DVEYKTL-- 348
Db 304 LANGVIDS-DGNVYTFDVKDKENKENTANITMPAYID----PENVTKGAVTLTGIGT 358
QY 349 SSVNKTIVVEYQRENNERTANLQSMFTNIDTKNHTVEQTIYINPLRSKATENV----- 402
Db 359 NTSKTYLIDYEKYGQPHNLSIKGTIDQIDKTNNTYQTIYVNE----SGDNVVLPAIT 413
QY 462 NINF---GNIDSPYIIKVIKSKYDPNKKDYTTTQQVTMQTTINEYTGEE--RTASYDNT 515
Db 473 KVEFPDDDDQITTPYIVVNGHIDPASTG-----DLALSTFYGYDSNFIWRSMSWDNE 526
QY 516 IASTSSGOGG-DLP--PEKTYKIGDY--VWEDVDKD-GIONTNDN 556
Db 527 VAFNNGSGDGDIDKPVVPEQPEGEIEPIPEDSDSPSGSDS 573

```

Search completed: July 30, 2002, 10:40:30
 Job time: 569 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)